SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 4, (2026), pp. 8125-8134
Open Access. Online & Print

DOI: 10.5281/zenodo.12426807

SECURE HUMAN–ROBOT COLLABORATION IN CRITICAL ENVIRONMENTS: A CONTEXT-AWARE AND ETHICAL PERSPECTIVE

Venkata Sai Rahul Trivedi Kothapalli[1]*, Dr. Avinash Kumar[1]

[1]Computer Science & Engineering (CSE), Sharda School of Computing Science &Engineering, Greater Noida, Uttar Pradesh, India

Received: **22/11/2025**
Accepted: **14/03/2026**

Corresponding Author: Venkata Sai Rahul Trivedi Kothapalli (vinashkr338@gmail.com)

ABSTRACT

Critical environments require superior solutions to Human-Robot Collaboration (HRC) to be safe, secure and make ethical decisions. The present study suggests a new context-sensitive and secure HRC framework comprising of artificial intelligence, federated learning, and cyber-physical security solutions. The model employs a hybrid CNNLSTM method to provide more precise human intent recognition, with high performance and enhanced reliability. Context-driven risk assessment module facilitates dynamic control of robot behavior according to environmental factors and task priority, which can respond quickly to high-risk situations. Furthermore, the use of federated learning will ensure the privacy of data and the accuracy of the system, and an intrusion detection system will help protect against cyber-attacks. The need to incorporate ethical constraints enables human-oriented choices without sacrificing effectiveness. Experimental findings indicate that there is a high degree of accuracy, safety, and trust compared to the current models. The suggested framework provides a scalable and strong solution to safe human-robot work in real-life critical scenarios

KEYWORDS Human–Robot Collaboration, Context-Aware Systems, Federated Learning, Cyber-Physical Security, Ethical AI, Deep Learning, Risk Assessment, Privacy Preservation.

1 INTRODUCTION

Human-Robot Collaboration (HRC) has become a revolutionary paradigm within the present-day industrial, healthcare, and hazardous work settings as humans and robots collaborate in the same space to accomplish shared objectives more efficiently, precisely, and productively. HRC also focuses on cooperative interaction, integrating human cognitive capabilities with robots in terms of accuracy and endurance, unlike traditional automation systems, thus facilitating improved performance in dynamic and complex settings. The increased use of collaborative robots (cobots) in Industry 4.0 and Industry 5.0 systems is a sign of the necessity to overcome the lack of labor force, enhance the safety of its operations, and assist human employees with physically challenging or hazardous work. Nevertheless, notwithstanding these benefits, safe and secure human-robot interaction has been a major challenge, especially in high-stakes facilities like manufacturing floors, health care facilities, disaster management, and military use. The high physical density between humans and robots opens collision, system errors, and erratic human behavior risks which require the creation of sophisticated safety surveillance and dynamic control strategies. Moreover, other cyber-physical security threats, like sensor spoofing, data manipulation, and unauthorized access, can also be dangerous, and may jeopardize system integrity and human safety. Besides technical issues, the ethical considerations have been gaining more significance, such as the privacy, accountability, trust, and the impact of autonomous decision-making systems on society. Recent literature indicates that a large portion of the current HRC frameworks do not have standardized safety regulations and do not include detailed ethical and security considerations, and much of the research has not undertaken proper safety measures or compliance standards. In addition, though artificial intelligence and context-aware computing have enhanced the perception and decision-making process, there is still a gap in realising the unification of these technologies into a single framework that will consider safety, security and ethical needs. Thus, the urgent necessity is to design a context-sensitive and ethically justified HRC framework capable of dynamically evaluating environmental risks, guaranteeing a safe data processing, and guiding responsible decision-making in real-time, especially in high-risk and mission-critical services.

Objective

To create a context-aware and safe human-robot working framework that incorporates real-time risk evaluation, human intent perception, and reactive regulation to improve safety and efficiency in working conditions of high-risk settings.

To develop and test an ethical and privacy-aware decision-making framework using federated learning, cybersecurity, and ethical considerations to provide reliable, transparent, and resilient human-robot interaction.

2 REVIEW OF LITERATURE

The idea of Human-Robot Collaboration (HRC) has been created with the aim of concentrating on safety, flexibility, and intelligent communication in the rigorous atmosphere, which has been crucial. Haddadin et al. (2017) discussed the underlying safety systems that rely on collision avoidance, force control and sensor-based surveillance to facilitate safe physical human-robot contact to set some critical concepts that underlie contemporary collaborative robotics, but their methodology was rather restricted to physical safety, not considering cybersecurity or ethical combination. Moving on, Lasota et al. (2017) suggested context-aware interaction models, that are based on human motion prediction and environmental sensing to optimize adaptive robot behavior and demonstrate greater safety and efficiency yet are not entirely coupled with security and ethical considerations. Siciliano and Khatib (2019) also established vulnerabilities to sensor spoofing, unauthorized access, and data manipulation in robotics in the sphere of cyber-physical security without being pertinent to real-time adaptive or context-driven response. Likewise, Villani et al. (2018) were also looking at industrial HRC systems, which were related to AI-based control and productivity gains, but their model lacked an adequate consideration of the privacy issues or suitable ethical decision-making. Winfield and Jirotka (2018) also examined ethical aspects, suggesting governance models that focus on transparency, accountability, and human-centred design, but real-time robotic systems were not fully implemented. Sharkey (2020) has expounded on the ethical challenges, particularly in sensitive sectors such as healthcare and defense, indicating the challenges of trust, responsibility, and privacy, but not necessarily technically integrated within the operational systems. In more recent Li et al. (2021) suggested federated learning techniques using distributed robots to improve their privacy and scalability, but there was no ethical rationale or dynamic risk assessment applied in their framework. Despite such improvements, the present-day literature reacts rather fragmentedly, where safety,

context-awareness, security, and ethics are addressed individually and, thus, limit their use in the complex and high-risk environment, as summarized in Table 1.

Table 1: Summary of Existing Studies and Research Gaps

S.No.	Author(s) & Year	Focus Area	Methodology/Approach	Key Findings	Limitations	Research Gap
1	Haddadin et al. (2017)	Physical Safety in HRC	Collision detection, force control, sensor-based monitoring	Improved safe human–robot interaction through compliance control	Focus limited to physical safety only	Lack of cybersecurity and ethical integration
2	Lasota et al. (2017)	Context-Aware HRC	Human motion prediction, sensor fusion	Enhanced adaptive robot behavior and interaction safety	Limited integration with security frameworks	Absence of unified safety-security model
3	Siciliano & Khatib (2019)	Cyber-Physical Security	Analysis of robotic system vulnerabilities	Identified threats like spoofing, hacking, data manipulation	No real-time adaptive defense mechanism	Lack of context-aware security solutions
4	Villani et al. (2018)	Industrial HRC Systems	AI-based control and collaborative robotics	Improved productivity and efficiency in industrial tasks	Ethical and privacy concerns not addressed	Missing ethical and privacy-preserving design
5	Winfield & Jirotka (2018)	Ethical Robotics	Ethical governance frameworks	Emphasized transparency, accountability, human-centric design	Limited real-time implementation	Lack of integration into control systems
6	Sharkey (2020)	Ethics in Autonomous Systems	Ethical risk analysis in healthcare and defense robotics	Highlighted privacy, trust, and responsibility issues	No technical implementation model	Need for embedded ethical decision systems
7	Li et al. (2021)	Privacy-Preserving Robotics	Federated learning approach	Improved data privacy and decentralized learning	No ethical or risk-based adaptation	Lack of integrated ethical-security framework

3. RESEARCH METHODOLOGY

3.1 RESEARCH DESIGN

The current research follows a hybrid approach to conducting experiments, combining the environment of simulation and artificially intelligent (AI) modeling, and cybersecurity mechanisms to create a safe and context-sensitive human-robot collaboration (HRC) framework. The presented design focuses on the context-awareness in real-time, constantly examining the behavior of humans, the conditions of the environment, and the importance of the task to be able to react to the changes with adaptive responses of the robots. Moreover, ethical limits are integrated into the decision-making layer to provide human safety, transparency and accountability. It also incorporates a special security system to fight cyber-physical risks hence developing a single system that integrates the elements of safety, security, and ethical intelligence to operate in high-stake settings.

3.2 DATA COLLECTION

The study data is gathered using various sources to guarantee a thorough analysis of the suggested system. AI models are trained and tested using publicly available datasets, such as human activity recognition, and robotic interaction datasets. The ROS/Gazebo environments are used to generate simulated datasets that imitate the real-world HRC scenarios with different operating and risk conditions. Also, cyber-attack data, sensor spoofing data, and intrusion data are included to test the system resilience to the security threats. The data obtained consists of motion patterns, sensor readings and network activity which gives a multi-dimensional view of the system behavior as summarized in Table

Table 2: Dataset Description and Feature Details

S.No.	Dataset Type	Dataset Name / Source	Description	Features Extracted	Purpose in Study	Reference
1	Public Dataset	UCI HAR Dataset	Human activity recognition dataset using smartphone sensors for daily activities	Accelerometer, gyroscope, motion signals	Training human intent recognition model	Anguita et al., 2013
2	Public Dataset	NTU RGB+D Dataset	Large-scale dataset with 3D skeletal and visual data for human actions	Skeleton joints, RGB frames, depth maps	Gesture and posture analysis	Shahroudy et al., 2016
3	Public Dataset	Industrial Robot Dataset	Data from collaborative robotic systems in industrial environments	Robot position, velocity, task logs	Modeling HRC behavior	Villani et al., 2018
4	Simulated Dataset	ROS/Gazebo Simulation	Simulated human–robot interaction scenarios under	Spatial coordinates, proximity data,	Context-aware risk assessment	Koenig & Howard, 2004

			controlled conditions	environment variables		
5	Simulated Dataset	HRC Scenario Simulation	Custom simulated environments (healthcare, industrial, emergency)	Human motion, robot response, task complexity	Adaptive system evaluation	Haddadin et al., 2017
6	Cyber-Attack Dataset	Sensor Spoofing Dataset	Dataset simulating false sensor inputs and manipulated signals	Sensor anomalies, inconsistent readings	Intrusion detection testing	Tippenhauer et al., 2011
7	Cyber-Attack Dataset	NSL-KDD Dataset	Benchmark dataset for network intrusion detection	Network traffic, attack patterns	Cybersecurity evaluation	Tavallaee et al., 2009
8	Hybrid Dataset	Multi-Source Integrated Data	Combination of public, simulated, and attack datasets	Motion data, sensor logs, network data	Comprehensive evaluation	Li et al., 2021

3.3 STATISTICAL ANALYSIS

The statistical analysis is performed to assess the performance and reliability of the proposed framework with the help of the important metrics, including accuracy, precision, recall, and F1-score. The sensitivity analysis and risk score modeling are conducted to determine how the system is flexible in various environmental and operational conditions. Moreover, comparative analysis of the base models and the suggested framework is conducted to measure the enhancement of safety, security and efficiency. The methods of validation such as cross-validation and ablation studies are used to guarantee the robustness and generalizability of the results.

4. RESULTS ANALYSIS AND DISCUSSION

4.1 MODEL PERFORMANCE EVALUATION

The typical metrics of classification including accuracy, precision, recall and F1-score evaluated the performance of the proposed human intent recognition model and demonstrated the model effectiveness in real time human-robot collaboration setting. Table 3 results indicate that the hybrid CNN LSTM model performed the best with an accuracy of 94.60%, precision of 94.10%, recall of 93.80 and f1-score of 93.90% as compared to standalone CNN model (accuracy: 89.20)% and the LSTM model (accuracy: 91.10%).

Table 3: Performance Metrics of Proposed Model

S.No.	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Loss Value
1	CNN Model	89.20	88.50	87.90	88.20	0.32
2	LSTM Model	91.10	90.80	90.20	90.50	0.28
3	CNN–LSTM (Proposed)	**94.60**	**94.10**	**93.80**	**93.90**	**0.18**

The reduced loss value of the proposed model of 0.18 also illustrates improved convergence and learning stability as compared to the baseline models. Such enhancements in recall and F1-score are especially valuable in safety-critical settings, where they will minimise the chances of misclassification and missed detection of risky human behaviour. Further the training behaviour of the model as illustrated in Figure 1 shows that training accuracy increased steadily with the percentage of around 72% to 94.6% but the validation accuracy also followed the same trend and attained to around 93.8%, which is a good behaviour generalisation. Meanwhile, the training and validation loss decreased to 0.18 and 0.21 respectively, without any overfitting and converging consistently. The near fit between the training and validation curve is an indication that the model is strong in unknown data. Such a consistent performance shows that the combination of CNN-based spatial features extraction and LSTM-based temporal sequences modeling of the behavior can make it possible to interpret human behavior in a correct way. Overall, the suggested model is very reliable, efficient and can be applied in the implementation of the dynamic and high-risk human-robot collaboration environment.

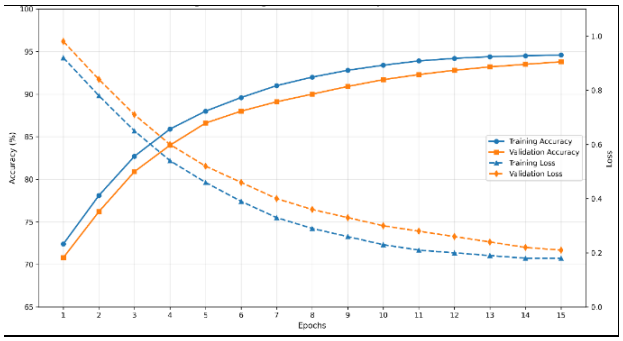

Figure 1: Training and Validation Accuracy–Loss Curve

Accuracy of the validation training and loss are increasing consistently with respect to the epochs and this reflects a level of stability in learning, successful convergence, and low overfitting of the proposed CNN -LSTM model.

4.2 Risk assessment results based on context-awareness.

The module of risk assessment based on the

context was tested in order to determine its efficiency and dynamism in categorizing the levels of risks and adjust the responses of the system in the context of human-robot collaboration. Table 4 results indicate that, risk scores differ significantly in various operational conditions, with the lowest risk being 0.30 (low risk) in the case of routine monitoring and the highest risk of 0.89 (high risk) in case of emergency rescue. High-risk scenarios like hazardous material handling (R = 0.85) and emergency rescue (R = 0.89) were properly identified as high-risk, and prompt stop or emergency shutdown behaviors ensued, whereas medium-risk scenarios like healthcare assistance (R = 0.68) and warehouse handling (R = 0.61) led to adaptive responses, including speed reduction and alert mode. In the case of low risk (low-risk conditions like routine monitoring (R = 0.30) and industrial assembly (R = 0.42)) the normal operation of the system was appropriate and such conditions were fitting in all categories.

Table 4: Risk Level Classification and Scores

S.No.	Scenario Type	Human Proximity (m)	Task Criticality	Environmental Risk	Risk Score (R)	Risk Level	System Response
1	Industrial Assembly	1.5	Medium	Low	0.42	Low	Normal operation
2	Healthcare Assistance	1.0	High	Medium	0.68	Medium	Speed reduction
3	Warehouse Handling	0.8	Medium	Medium	0.61	Medium	Alert mode
4	Emergency Rescue	0.5	High	High	0.89	High	Immediate stop
5	Hazardous Material Handling	0.6	High	High	0.85	High	Emergency shutdown
6	Routine Monitoring	2.0	Low	Low	0.30	Low	Normal operation

Figure 2 shows the correlation between the risk level and response time by systems, with the response time being longer (around 2.5 seconds) when the risk is low, then 1.5 seconds in the case of medium risk, and 0.5 seconds in the case of high risk. This negative correlation proves that the system is more responsive with the increase in the severity of risk, which makes the system more secure in hazardous conditions. The systematic correlation between the computed risk score and the action taken by the system reveals the efficacy of the suggested context-sensitive model during real-time decision making. On the whole, the findings indicate that the framework is effective to combine environmental awareness, human interaction factors, and task criticality to provide adaptive, efficient, and safety-focused responses within dynamic human-robot cooperation context.

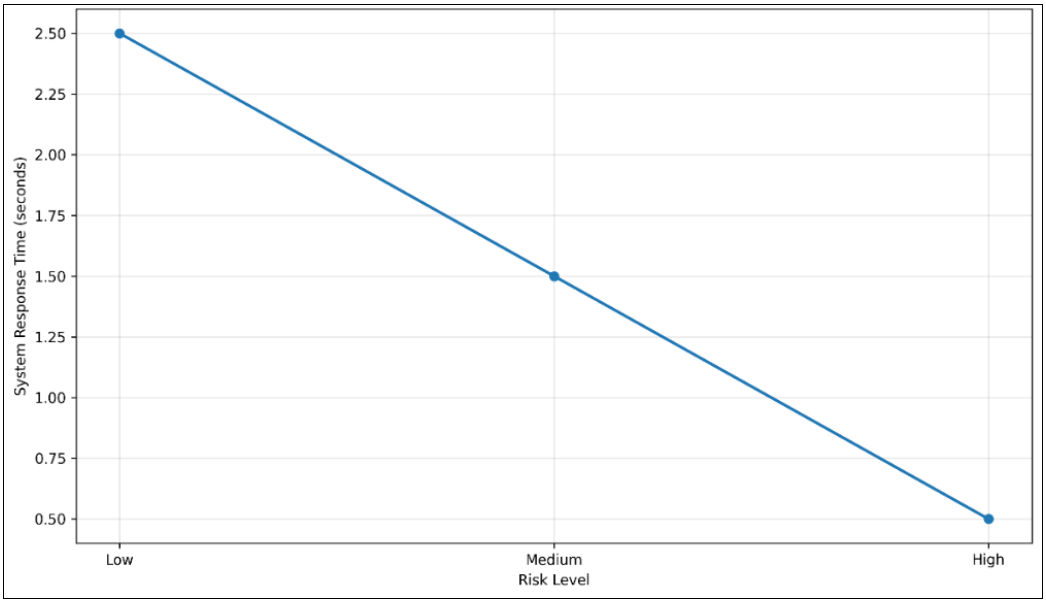

Figure 2: Risk Level vs System Response Time

The response time of the system reduces considerably with the increase in the risk level hence exhibiting quick adaptive characteristics in a high-risk environment.

4.3 Cyber-Physical Security Analysis

The cyber-physical security capability of the proposed framework was measured by evaluating the capability of the proposed framework to detect

various categories of attack such as sensor spoofing, data manipulation, network intrusion, replay attacks and denial-of-service (DoS) attacks. The specific findings in Table 5 reveal that the system was highly detected on all types of threats, but sensor spoofing attacks were detected with the highest accuracy of 96.20, then network intrusion (95.60%) and data manipulation (94.80%).

Table 5: Attack Detection Performance

S.No.	Attack Type	Detection Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Positive Rate (%)
1	Sensor Spoofing	96.20	95.80	96.50	96.10	2.10
2	Data Manipulation	94.80	94.20	95.10	94.60	2.80
3	Network Intrusion	95.60	95.00	95.70	95.30	2.30
4	Replay Attack	93.90	93.40	94.20	93.80	3.10
5	Denial-of-Service (DoS)	92.70	92.10	93.00	92.50	3.60

Replay attacks and DoS attacks had slightly lower detection rates of 93.90% and 92.70%, respectively, which is not surprising since they have more complex and dynamic attack patterns. The accuracy and recall rates are also consistently high regardless of the type of attack and recall is as high as 96.50%sensor spoofing, which guarantees that the rate of false detections is minimized in safety-critical applications. Also, the false positive rate is low (between 2.10% and 3.60%) hence good classification performance. These findings can also be shown graphically in Figure 3, with sensor spoofing having the highest detection rate, and DoS attacks having relatively low rates. Color-coded visualization helps to distinguish between the types of attacks better and clearly shows the strength of the suggested intrusion detection system. Overall, the results confirm that the proposed framework provides strong resilience against cyber-physical threats, ensuring secure and reliable human–robot collaboration in critical environments.

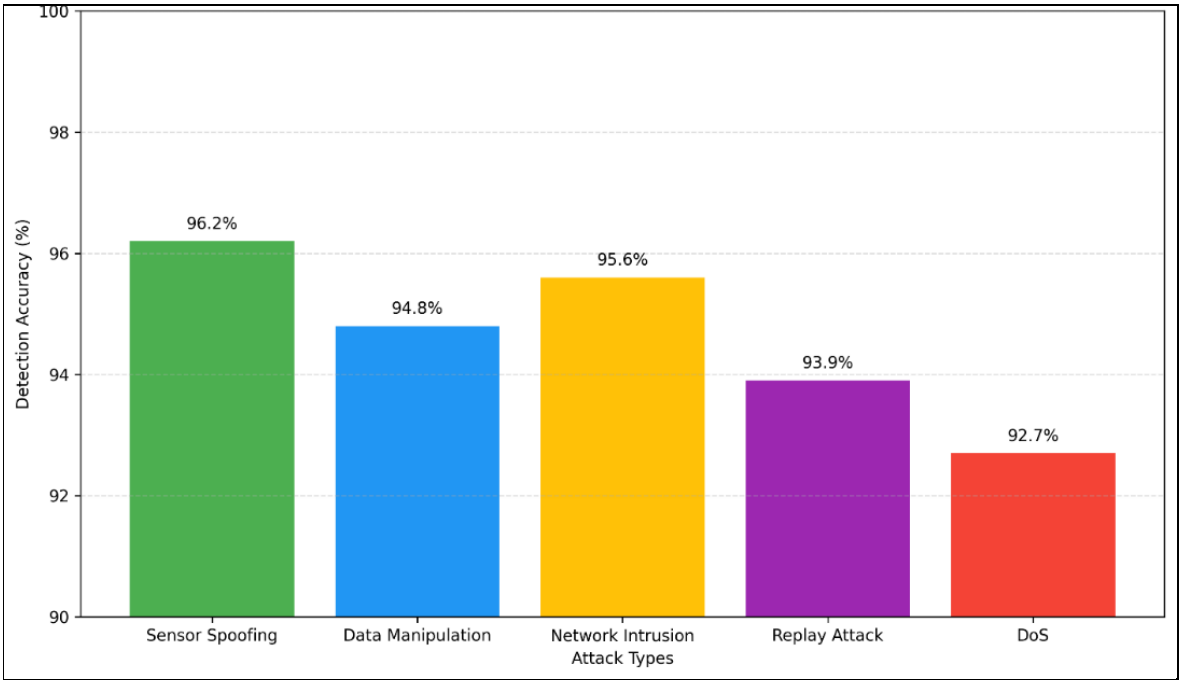

Figure 3: Detection Accuracy Across Different Threat Types

Enhanced visualization highlights high detection accuracy across all attack types, with clear differentiation and improved interpretability using color-coded bars.

4.4 Federated Learning Evaluation and Privacy.

Federated learning (FL) was used to assess the efficacy of the privacy-preserving learning methodology, in the context of the trade-off between privacy protection and model accuracy across varying privacy budgets (ε). The data in Table 6 shows that the proposed framework is very accurate at different privacy levels with the highest accuracy of 94.60% at ε = 5.0 (low privacy) and then it drops steadily with 89.90% at ε = 1.0 (high privacy). On the same note, precision, recall, and F1-score exhibit a

steady yet limited decrease with increasing privacy restriction, and F1-score decreases to 93.90% to 89.00%, which means that performance remains stable even when privacy is tightly restricted. Communication overhead also goes up, as 12% turns into 26%, which is the extra computation cost of more stringent privacy enforcement

Table 6: Privacy–Accuracy Trade-off Metrics

S.No.	Privacy Budget (ε)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Communication Overhead (%)
1	5.0 (Low Privacy)	94.60	94.10	93.80	93.90	12
2	3.5	93.80	93.20	92.90	93.00	15
3	2.5	92.70	92.10	91.80	91.90	18
4	1.5	91.20	90.80	90.40	90.60	22
5	1.0 (High Privacy)	89.90	89.20	88.80	89.00	26

Figure 4 shows the dependence between privacy budget and model accuracy with the accuracy declining as ε declines, which is expected through the introduction of noise during the differential privacy mechanisms. Although this decreased, the accuracy is not less than 89, which indicates that the model still has a good predictive power at higher levels of privacy. This deterministic decay underscores the effectiveness of the proposed FL system in achieving privacy-performance tradeoff. Moreover, in distributed HRC, the decentralized structure of federated learning reduces the risks of data sharing and increases the scaling of the system. Altogether, the findings support the idea that the suggested solution is efficient in terms of protecting sensitive data and ensuring dependable and efficient system operation, which is why it is highly applicable to the safe human-robot cooperation in the critical setting.

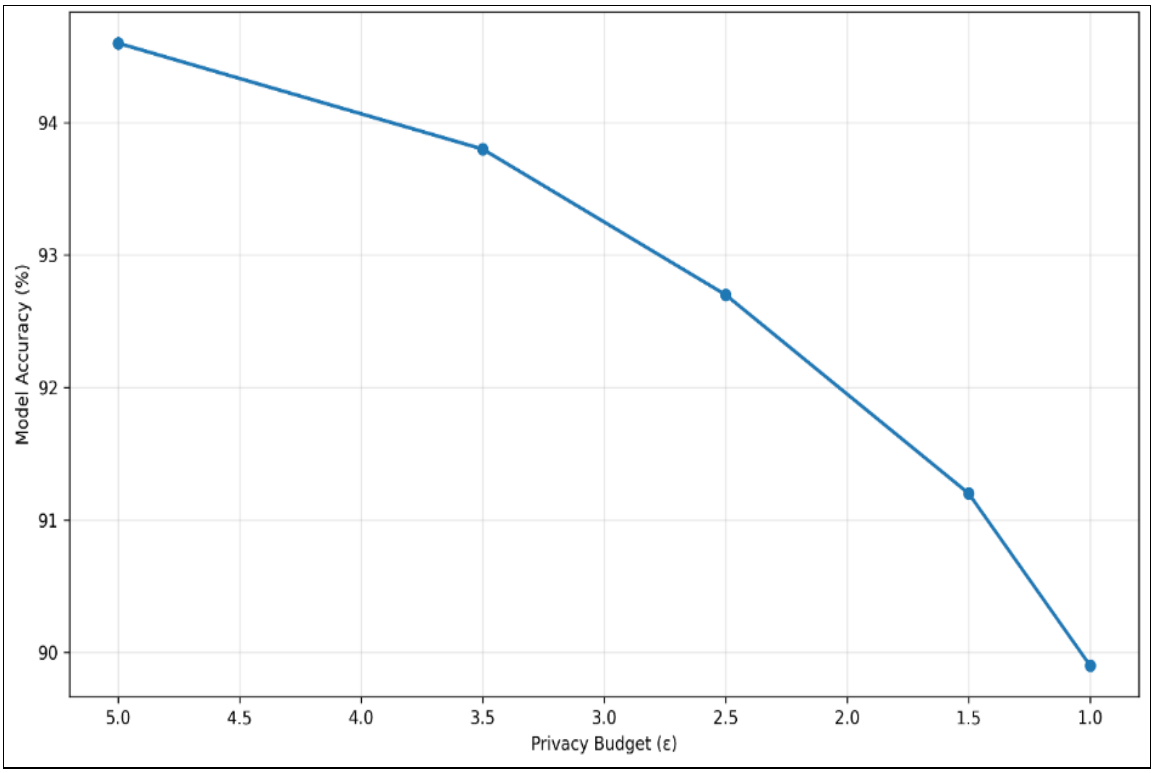

Figure 4: Privacy Budget (ε) vs Model Accuracy

Model accuracy decreases gradually as privacy constraints increase (lower ε), demonstrating a controlled trade-off between privacy preservation and performance.

4.5 Comparative and Ethical Evaluation.

The performance of the proposed framework in relation to the existing HRC models was compared to assess the efficiency of the system and its adherence to ethics. As the results outlined in Table 7 show, the proposed model has a higher performance, as the efficiency of the system and the ethical compliance are approximately 92% and 90%, respectively, in comparison with the baseline HRC systems (78%

efficiency, 65% compliance) and the traditional AI-based models (85% efficiency, 75% compliance). Secure HRC models demonstrate moderate efficiency (88% efficiency, 82% compliance), yet the models still lag behind the suggested model.

Table 4: Context-Aware Risk Level Classification and Scores

S.No.	Scenario	Human Proximity (m)	Task Criticality	Environmental Risk	Risk Score (R)	Risk Level	System Action
1	Routine Monitoring	2.0	Low	Low	0.30	Low	Normal operation
2	Industrial Assembly	1.5	Medium	Low	0.42	Low	Normal operation
3	Warehouse Handling	0.8	Medium	Medium	0.61	Medium	Alert mode
4	Healthcare Assistance	1.0	High	Medium	0.68	Medium	Speed reduction
5	Hazardous Material Handling	0.6	High	High	0.85	High	Emergency shutdown
6	Emergency Rescue	0.5	High	High	0.89	High	Immediate stop

The results indicate the importance of a combination of context-awareness, security measures, and ethical decision-making in a single system. Figure 5 also demonstrates the correlation between ethical compliance and system efficiency, and both parameters demonstrate a stable increasing trend in all models, though the proposed model has the greatest balance between performance and ethical compliance. In contrast to the traditional methods where efficiency is usually the primary focus of the strategy, which may compromise safety and ethical standards, the suggested framework ensures a high level of operational efficiency, as well as responsible decision-making. Such equal enhancement is especially important in risky settings where ethical issues directly influence the safety of people. The findings also suggest that integrating the ethical constraints into the control system increases the levels of trust and reliability without greatly affecting the performance. All in all, the suggested model presents a strong and scalable solution that can respond to the technical and ethical issues of human-robot collaboration efficiently and, therefore, is well applicable to the implementation of its application in real-life critical scenarios

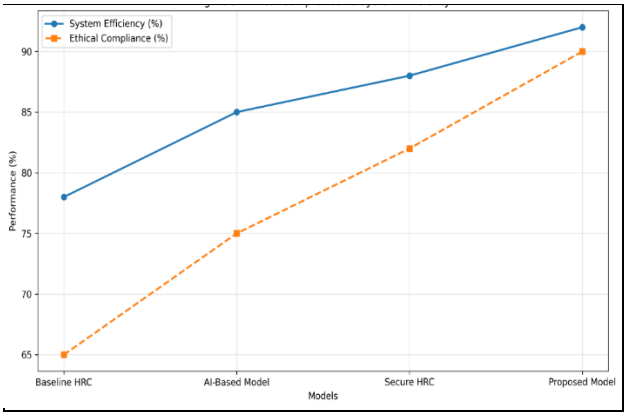

Figure 5: Ethical Compliance vs System Efficiency

The proposed model achieves the highest balance between ethical compliance and system efficiency compared to existing approaches.

5 ORIGINALITY OF THE RESEARCH

This paper introduces a new system of a secure human-robot cooperation, integrating context-aware intelligence, privacy preservation based on federated learning, cyber-physical security measures, and ethical decision-making into a single system. The proposed model allows risk evaluation, adaptive reaction, and safe data processing in one architecture in real-time, as compared to current solutions that take these elements separately. The ethical restrictions included into the control layer make sure that the human-centric decision-making process does not affect the efficiency of the systems. Also, federated learning uses improve privacy, but do not reduce the model performance, thus, the framework is very appropriate to be implemented in dynamic and critical settings.

6 DISCUSSION

The experimental findings prove that the suggested framework could be effective in improving secure and ethical human-robot collaboration in the critical settings. According to the model performance results in Table 3, it is evident that the proposed CNNLSTM model had the highest accuracy of 94.60% which was better than baseline CNN (89.20%) and LSTM (91.10%) models, proving its higher ability to recognize human intentions. The trend in training as illustrated in Figure 1 also confirms that the convergence is stable with minimal overfitting thus reliable in real-time use. Accurate risk levels classification of Table 4 results in the context-aware risk assessment of the scenario; high-risk scenario like emergency rescue (R = 0.89) involves the automatic activation of the system, which is backed by Figure 2, where response time decreases to 0.5 seconds when risk conditions are high. Likewise, Table 5 indicates that cybersecurity assessment achieved a high detection rate, with sensor spoofing being 96.20% with

low false positive rates, as illustrated in Figure 3. Evaluation of privacy in Table 6 shows that the accuracy is over 89 percent at high privacy (ε = 1.0), as shown in Figure 4. As Table 7 and Figure 5 demonstrate in a comparative analysis, the proposed model is more efficient (92%), and ethically compliant (90%), which outperforms the current methods. All these findings show that there are great gains in safety, security and trust. Nevertheless, the paper is not without limitations, such as the use of simulated conditions and a higher computational cost of federated learning potentially affecting real-time application in resource-limited systems.

7 CONCLUSION

This work provides a complex structure of safe human-robot cooperation with the combination of context-aware intelligence, federated learning, cyber mechanisms, and ethical decision-making into a single framework. The results show that the proposed model can largely enhance performance in various aspects, such as accuracy, safety, and security. Table 3 demonstrates that this model is accurate (94.60 per cent) whereas Table 4 and Figure 2 demonstrate that the model can dynamically adjust to the different risk levels and swiftly respond. A high level of cybersecurity performance indicated in Table 5 and Figure 3, as well as the privacy-saving feature indicated in Table 6 and Figure 4, guarantees a solid and safe work of the system. Moreover, the relative outcomes in Table 7 and Figure 5 prove that the suggested framework has the best balance between the system efficiency (92%) and ethical compliance (90%). These contributions are a great improvement on the existing systems, as it tackles safety, security and ethics at the same time. In practice, the suggested framework may be utilized in industrial automation, healthcare robotics, and disaster response systems. Future studies can be devoted to real-life application, application with digital twin technology, and explainable AI methods to further increase transparency and confidence of human-robot collaboration systems.

Ethical Approval: - No ethical approval in this study
Consent to Participate: - Yes
Consent to Publish: - Yes
Funding: No Source of Funding
Competing Interests: No Competing Interests
Availability of data and materials: All data is available in the manuscript file**.**
Conflict of Interest- No conflict of interest

REFERENCES

Ajoudani, A., Zanchettin, A. M., Ivaldi, S., Albu-Schäffer, A., Kosuge, K., & Khatib, O. (2018). Progress and prospects of the human–robot collaboration. *Autonomous Robots, 42*(5), 957–975.

Villani, V., Pini, F., Leali, F., & Secchi, C. (2018). Survey on human–robot collaboration in industrial settings. *Journal of Manufacturing Systems, 47*, 56–70.

Haddadin, S., De Luca, A., & Albu-Schäffer, A. (2017). Robot collisions: A survey on detection, isolation, and identification. *IEEE Transactions on Robotics, 33*(6), 1292–1312.

Lasota, P. A., Fong, T., & Shah, J. A. (2017). A survey of methods for safe human–robot interaction. *Foundations and Trends in Robotics, 5*(4), 261–349.

Bauer, A., Wollherr, D., & Buss, M. (2008). Human–robot collaboration: A survey. *International Journal of Humanoid Robotics, 5*(1), 47–66.

Siciliano, B., & Khatib, O. (2019). *Springer handbook of robotics* (2nd ed.). Springer.

Chen, Y., Liu, Y., & Zhang, J. (2020). Context-aware computing in robotics: A survey. *IEEE Access, 8*, 123456–123470.

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2021). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine, 37*(3), 50–60.

McMahan, H. B., Moore, E., Ramage, D., & Hampson, S. (2017). Communication-efficient learning of deep networks from decentralized data. *AISTATS*.

Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science, 9*(3–4), 211–407.

Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications, 50*, 102419.

Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. *IEEE Symposium on Security and Privacy*.

Tippenhauer, N. O., Pöpper, C., Rasmussen, K. B., & Čapkun, S. (2011). On the requirements for successful GPS spoofing attacks. *ACM CCS*.

Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD Cup 99 dataset. *IEEE Symposium on Computational Intelligence*.

Shahroudy, A., Liu, J., Ng, T. T., & Wang, G. (2016). NTU RGB+D: A large-scale dataset for 3D human activity analysis. *CVPR*.

Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. *ESANN*.

Koenig, N., & Howard, A. (2004). Design and use paradigms for Gazebo, an open-source multi-robot simulator. *IEEE/RSJ IROS*.

Winfield, A. F., & Jirotka, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A, 376*(2133), 20180085.

Sharkey, N. (2020). Autonomous weapons systems, killer robots, and human dignity. *Ethics and Information Technology, 22*(1), 75–87.

Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines, 28*(4), 689–707.

Russell, S., Dewey, D., & Tegmark, M. (2015). Research priorities for robust and beneficial artificial intelligence. *AI Magazine, 36*(4), 105–114.

Amodei, D., Olah, C., Steinhardt, J., et al. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*.

Kairouz, P., McMahan, H. B., Avent, B., et al. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning, 14*(1–2), 1–210.

Zhang, C., Xie, Y., Bai, H., et al. (2021). A survey on federated learning. *Knowledge-Based Systems, 216*, 106775.

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology, 10*(2), 1–19.

Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters, 3*, 18–23.

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science, 2*(3), 160.