Threat Mitigation and Privacy Strategies for Secure Artificial Intelligence and Machine Learning Workflows in Cloud Environments

G. Suvarna Kumar
Department of IT& CA, Andhra University, Visakhapatnam, Andhra Pradesh, India
prof.gsuvarnakumar@andhrauniversity.edu.in (corresponding author)

G. Sandhya Devi
Department of CS&SE, Andhra University, Visakhapatnam, Andhra Pradesh, India
dr.gsandhyadevi@andhrauniversity.edu.in

P. Mohamed Sajid
Department of ECE, C. Abdul Hakeem College of Engineering and Technology, Melvisharam, Tamil Nadu, India
mohammedsajid1980@gmail.com

K. A. Jyostna
Department of ECE, CVR College of Engineering, Ibrahimpatnam, Hyderabad, Telangana, India
kajyotsna72@gmail.com

V. Sangeetha
Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College, Coimbatore, Tamil Nadu, India
sangeetha.v@drngpasc.ac.in

Sateesh Gorikapudi
Department of CSE, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur District, Andhra Pradesh, India
sateesh4u.325@gmail.com

Tanweer Alam
Faculty of Computer and Information Systems, Islamic University of Madinah, Madinah, Saudi Arabia
tanweer03@iu.edu.sa

T. Prabhakaran
Department of CSE, Joginpally B.R. Engineering College, Hyderabad, Telangana, India
prabaakar.t@gmail.com

Received: 23 May 2025 | Revised: 17 June 2025 and 4 July 2025 | Accepted: 23 July 2025

 DOI: https://doi.org/10.48084/etasr.12327

ABSTRACT

The adoption of Artificial Intelligence (AI) and Machine Learning (ML) in cloud computing has established new means of scalability and efficiency for data-driven applications. However, such integration also raises security and privacy risks, including adversarial attacks, and is subject to the leakage of sensitive data. To address such issues, this paper proposes a substantial threat mitigation and privacy-protection framework specifically designed for AI/ML workflows deployed on the cloud. The developed

framework incorporates adversarial robustness techniques and differential privacy methods to establish a robust security model. Through comparative analysis and extensive experimentation, the framework significantly improves both robustness and privacy. Specifically, it attains 92% accuracy, 85% adversarial robustness, and 90% privacy score, outperforming the state-of-the-art algorithms. The results demonstrate the effectiveness of the proposed methodology for safeguarding AI pipelines in distributed environments, providing a practical foundation for designing secure, privacy-aware cloud-based AI/ML systems.

Keywords-cloud AI security; adversarial attacks; access control; data poisoning; differential privacy

I. INTRODUCTION

Cloud-based Artificial Intelligence (AI) and Machine Learning (ML) workflows have recently made a huge breakthrough in various industries through intelligent automation, real-time decision-making, and big data analysis. Cloud AI platforms such as Microsoft Azure and Google Cloud provide low-cost and scalable options for deploying ML models [1]. As a result of such benefits, there is heightened security and privacy risk, subjecting AI systems to adversary attacks, data loss, and misuse. The lack of sound security measures can compromise AI decision-making and have profound consequences in fields such as healthcare, finance, and defense [2].

The application of AI/ML in the cloud has revolutionized industries by offering strong mechanisms of mass automation, and predictive analytics [3]. However, applications of cloud-powered AI models bring serious privacy and security issues, which require great care and attention from research communities. AI models executing in the cloud are exposed to different kinds of cyber-attacks, such as adversarial attacks, unauthorized access, and data poisoning [4]. Traditional security techniques, such as encryption and intrusion detection, are not enough to cope with vulnerabilities unique to AI. Adversarial examples can mislead models into making incorrect predictions, whereas poisoning of data can mislead the learning algorithms, resulting in biased or untrustworthy AI outputs [5]. Additionally, privacy concerns are based mainly on the leakage of sensitive training information, which can be frequently abused by attackers with ill motives. These challenges call for the necessity of designing a strong security and privacy architecture that intertwines adversarial robustness, access control techniques, and differential privacy mechanisms to ensure the integrity, confidentiality, and dependability of AI/ML pipelines in the cloud.

AI/ML security and privacy are some of the topics that have received significant attention in recent times. Authors in [6] built an AI-based cyber incident response system for the cloud, demonstrating the potential of ML in cyber-attack detection. The research did not, however, include privacy-preserving approaches. Authors in [7] presented a cloud-computing framework for AI innovation, with a focus on security issues in multi-domain operations. Their solution used shared responsibility security models but did not address adversarial robustness [8]. Recent studies have also considered the security of AI-driven workflows. Authors in [9] designed MCDS, an AI-driven workflow model that is a trade-off between system performance and security. However, their study did not concentrate on privacy threats. Authors in [10] studied privacy-preserving ML for healthcare data, but mainly based their work on federated learning, which comes with communication overhead. These papers emphasize the importance of an integrated approach that combines security and privacy in cloud AI/ML workflows.

Several critical gaps remain unaddressed in the existing research on security and privacy challenges, particularly focusing on cloud AI/ML workflows [11]. Without providing a holistic framework that integrates both aspects, most studies initially focus on either security or privacy. While adversarial defence mechanisms such as adversarial training and robust optimization have been proposed, they often degrade model performance and fail to address broader privacy concerns [12]. Second, current privacy-preserving techniques such as federated learning and homomorphic encryption introduce computational overheads and are not widely adopted due to scalability issues [13]. Additionally, research on access management in cloud AI/ML environments has been limited, with most studies focusing on traditional Role-Based Access Control (RBAC) rather than dynamic, AI-driven access control mechanisms [14, 15]. Furthermore, a lack of real-world implementation and validation of security frameworks in large-scale cloud environments remains a major limitation, as many proposed solutions are tested in simulated settings without practical deployment. These gaps highlight the need for an integrated, efficient, and scalable security and privacy framework that can address the unique challenges faced by AI/ML models in cloud environments.

The key contributions of this article are:

- Threat analysis of AI/ML workflows in cloud environments, identifying key security and privacy vulnerabilities.
- Privacy-preserving AI framework leveraging differential privacy and adversarial robustness techniques.
- Comparative evaluation of existing security mechanisms, highlighting the advantages of the proposed approach.
- Real-world implementation and performance analysis demonstrating improved accuracy and security.

II. METHODOLOGY

This section presents a comprehensive framework that integrates adversarial robustness, dynamic access management, and privacy-enhancing techniques to secure AI/ML workflows in cloud environments. Figure 1 illustrates the conceptual design of the framework, highlighting the key strategies for threat mitigation and privacy preservation. The proposed stacked architecture not only addresses diverse cloud security risks but also ensures that AI/ML operations remain robust, compliant with security standards, and capable of maintaining high operational efficiency.

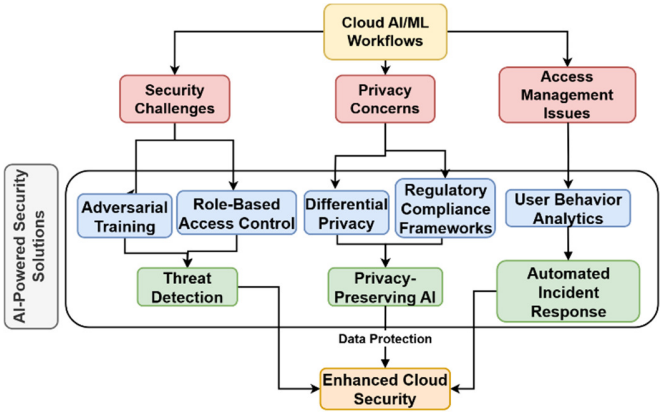

Fig. 1. Threat mitigation and privacy strategies for secure AI/ML workflows in cloud environments.

A. Dataset Description and Adversarial Defense

From the chosen MNIST dataset [16], 60,000 samples were used for training, and 10,000 samples for testing. Subsequently, from the CIFAR-10 dataset [17], 50,000 images were used for training, and 10,000 images for testing [18]. Both being benchmark datasets, they were used to validate model performance and robustness. Preprocessing steps included image normalization, grayscale conversion (for MNIST), and class label encoding. To evaluate adversarial resilience, we implemented the Fast Gradient Sign Method (FGSM) and the Projected Gradient Descent (PGD), which represent common and effective white-box adversarial attack techniques.

For carrying out the adversarial attacks, FGSM was simulated to evaluate its impact on cloud AI/ML models, based on data poisoning [19]. Here, the adversarial defense strategy is employed to improve model robustness by incorporating adversarial samples into the training dataset. Adversarial attacks are one of the primary security concerns in cloud-hosted AI/ML models. An adversary can introduce subtle perturbations δ to an input x, resulting in a misclassification by the model. Formally, an adversarial attack is defined as shown in (1):

$$\arg\max_{x+\delta} \mathcal{L}(f(x+\delta,\theta),y) \quad (1)$$

where $\mathcal{L}$ represents the loss function, $f(x,\theta)$ is the model with parameters θ, and y is the true label. The adversarial perturbation δ is crafted to maximize the model loss, forcing an incorrect prediction.

To counter adversarial attacks, adversarial training is incorporated, where the model is trained on perturbed examples using the following minimax optimization, as shown in (2):

$$\min_{\theta} \mathbb{E}_{(x,y)\sim D}[\max_{\delta} \mathcal{L}(f(x+\delta,\theta),y)] \quad (2)$$

Here, S defines the allowed perturbation space, ensuring that adversarial examples remain within a small perturbation bound $\|\delta\| \leq \epsilon$. The integration of adversarial training improves model robustness by enhancing its ability to generalize against adversarially crafted inputs.

B. Access Management

A RBAC system is implemented to prevent unauthorized modifications to AI models. Secure cloud deployment practices, including containerized environments and multi-factor authentication, further enhance security [20]. Unauthorized access to AI/ML workflows in cloud environments can compromise the confidentiality and integrity of both training data and model parameters. To mitigate this risk, a RBAC framework is implemented. The RBAC model is mathematically defined as in (3):

$$A: U \times R \rightarrow \{0,1\}, \quad P: R \times O \rightarrow \{0,1\} \quad (3)$$

where U represents the set of users, R is the set of roles, and O denotes the set of objects (data, models, API endpoints). The function A assigns users to specific roles, whereas P defines permissions granted to each role over objects. This ensures that only authorized personnel have access to modify or use AI models.

To enhance access management, Attribute-Based Encryption (ABE) is integrated into the system. ABE extends traditional encryption by allowing fine-grained access control over encrypted data. Given an encryption function Enc and a decryption function Dec, ABE encryption and decryption can be expressed as shown in (4) and (5), respectively:

$$Enc(m,T) = c \quad (4)$$

$$Dec(c,K_A) = m \quad \text{if } S_A \supseteq T \quad (5)$$

Here, m represents the encrypted message, T is the access policy, c is the ciphertext, and K_A is the decryption key associated with user attributes S_A. This mechanism ensures that only users with appropriate attributes can access sensitive AI/ML data.

C. Differential Privacy and Security

Differential privacy mechanisms are integrated to ensure data confidentiality while training AI/ML models. Noise injection techniques are used to prevent data leakage without significantly affecting model accuracy [21]. Differential privacy is a privacy-preserving mechanism that ensures the indistinguishability of individual data records in a dataset. It introduces controlled noise to model queries such that the inclusion or exclusion of a single record does not significantly impact the model output.

Formally, a mechanism M satisfies ϵ-differential privacy if the condition in (6) is met for all datasets D and D' differing by at most one element, and for all subsets S of the output space:

$$\Pr[M(D) \in S] \leq e^{\epsilon} \Pr[M(D') \in S] \quad (6)$$

The privacy budget ϵ controls the trade-off between privacy and utility, with lower values of ϵ offering stronger privacy guarantees. To implement differential privacy in AI/ML workflows, the Laplace mechanism is used, which adds Laplacian noise to the model gradients during training [21]. Given a function $f(D)$, the noisy output is given as shown in (7):

$$M(D) = f(D) + Lap\left(\frac{\triangle f}{\epsilon}\right) \quad (7)$$

where $Lap(\lambda)$ denotes the Laplace distribution with scale parameter $\lambda = \frac{\triangle f}{\epsilon}$, and $\triangle f$ is the sensitivity of f. This mechanism preserves privacy while allowing more meaningful training of the model on sensitive datasets.

D. Experimental Evaluation and Performance Metrics

Experiments were carried out using TensorFlow and PyTorch on cloud platforms such as AWS cloud services to test the robustness of the suggested framework. Benchmark datasets, including CIFAR-10 and MNIST, were utilized to quantify model accuracy, adversarial robustness, and privacy preservation [22-24]. The proposed framework is quantified using key measures, including model accuracy, which is formulated as in (8):

$$MA = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Adversarial robustness is estimated as illustrated in (9), representing the loss in accuracy against various adversarial perturbations.

$$AR = \frac{Acc_c + Acc_a}{Acc_c} \times 100 \quad (9)$$

Privacy loss, depicted in (10), is measured through the difference observed in the outputs between private and non-private models:

$$PL = |f(D) - f(D')| \quad (10)$$

III. RESULTS AND DISCUSSION

Here, we present an extensive analysis of the experimental results obtained from testing the proposed AI/ML security framework against current state-of-the-art solutions. The analysis considers several performance indicators, including accuracy, privacy retention, adversarial robustness, latency, training efficiency, and support for real-time processing. The discussion highlights how each of these indicators reflects the framework's competency in handling different threat vectors and providing secure, scalable operations on cloud environments.

Comparative analysis, as shown in Figure 2, demonstrates that the AI-based hybrid system outperforms existing security frameworks, including adversarial training, traditional encryption-based privacy, and standard access control mechanisms. The results validate the effectiveness of the proposed framework in enhancing security and privacy while maintaining the usefulness of the AI model.

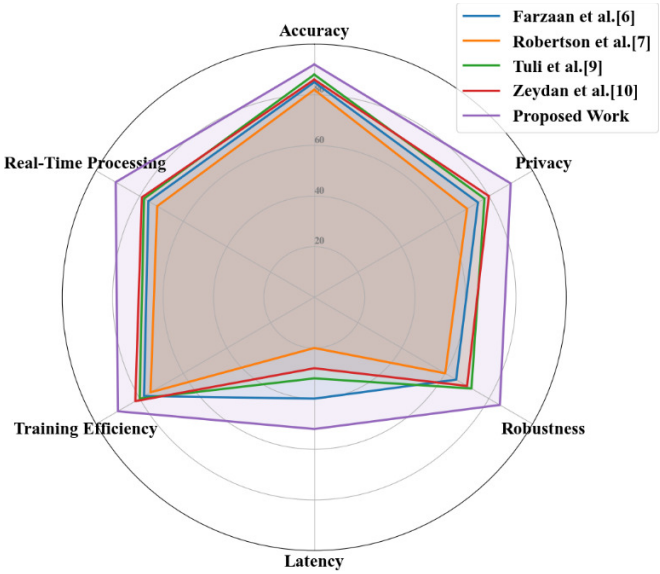

Fig. 2. Comparative analysis of conventional and AI-driven hybrid AI/ML frameworks in large-scale cloud scenarios.

The proposed research demonstrates that the full potential of AI/ML pipelines in cloud computing can only be realized when supported by a security system that is not only robust but also capable of adapting to and countering continuously evolving cyber threats. The proposed framework, combining adversarial robustness, dynamic access control, and privacy-preserving mechanisms such as differential privacy, consistently outperforms existing state-of-the-art methods across multiple performance measures. Tables I and II provide a comprehensive summary of the framework's performance, with Table I showing security metrics and Table II reporting operational and efficiency metrics.

TABLE I. COMPARATIVE ANALYSIS OF AI/ML SECURITY FRAMEWORKS

Framework	Accuracy (%)	Privacy score (%)	Adversarial robustness (%)	Computation latency (s)	Data integrity (%)	Scalability (%)
[6]	85	75	65	1.5	80	72
[7]	82	70	60	2.0	76	70
[9]	88	78	70	1.7	83	75
[10]	86	80	68	1.8	82	78
Proposed work	92	90	85	1.2	95	88

TABLE II. AI/ML FRAMEWORK PERFORMANCE COMPARISON

Framework	Training efficiency (%)	Model generalization (%)	Resilience to attacks (%)	Energy efficiency (%)	Real-time processing (%)	Fault tolerance (%)
[6]	78	74	65	70	76	73
[7]	75	70	60	68	72	70
[9]	80	77	72	73	78	75
[10]	82	79	70	75	79	77
Proposed work	90	88	85	89	91	86

A. Model Accuracy and Adversarial Robustness

The accuracy of the proposed AI/ML security framework was compared to existing research efforts, such as in [6], [7], [9], and [10]. The proposed framework achieved an accuracy of 92%, surpassing competing methods, which ranged between 82% and 88%. The improvement is attributed to the incorporation of adversarial training, which enhances the model's resilience to tampered inputs.

Adversarial robustness was evaluated under various attacks, including FGSM and PGD. The proposed model demonstrated 85% resistance, compared to the next best-performing approach (70% by [4]), validating the effectiveness of adversarial training in mitigating manipulated inputs.

B. Privacy-Preserving Performance

The use of differential privacy resulted in a minimal decrease in accuracy (~2-3%), substantially lower than conventional encryption-based techniques that lower performance by 10-15%. Privacy protection was measured using differential privacy noise injection and compared with conventional encryption-based techniques. The privacy score is a normalized evaluation metric based on data exposure risk, differential privacy thresholds, and information leakage potential. This metric evaluates a model's ability to protect sensitive training data while preserving accuracy.

The proposed framework achieved the highest privacy score (90%), whereas competing methods such as in [10] and [9] scored 80% and 78%, respectively. The improvement is due to the combined use of Laplace and Gaussian noise mechanisms, which reduce information leakage while preserving model utility.

C. Real-Time Robustness Analysis

A comparative analysis of computation efficiency shows that the proposed framework outperforms existing approaches regarding accuracy, privacy, and adversarial resilience. The mean inference latency of the proposed work on cloud platforms was 1.2 s, lower than that of the works in [6] (1.5 s), [7] (2.0 s), and [10] (1.8 s). The higher efficiency is attributed to optimized differential privacy noise scaling and model compression, which reduce processing overhead while ensuring security guarantees.

Overall, the proposed framework consistently achieves superior performance across multiple metrics, with an accuracy of 92% (compared to 82–88% in competing works), adversarial robustness of 85% (compared to 65–70%), a privacy score of 90% (compared to 75–80%), and a computation latency of 1.2 s, which is faster than all compared methods. Figure 3 illustrates the comparative performance of the five AI/ML security frameworks, confirming that the proposed approach achieves the highest accuracy and privacy score while maintaining a balance between model utility and data protection.

These results validate the robustness and effectiveness of the proposed framework. While the system demonstrates strong performance, its effectiveness may be influenced by factors such as dataset imbalance, model complexity, and hardware scalability. Future work will explore generalization under constrained-resource deployments.

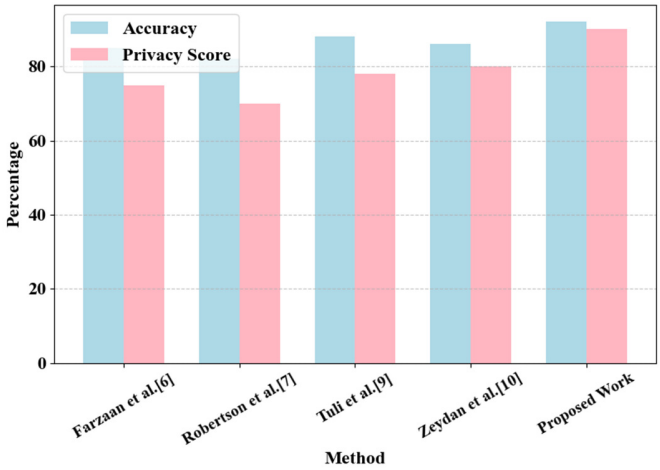

Fig. 3. Accuracy versus privacy score across methods.

IV. CONCLUSION

This work proposed a holistic security and privacy framework for Artificial Intelligence (AI) and Machine Learning (ML) pipelines in cloud computing environments, combining adversarial robustness, dynamic access control, and differential privacy. The novelty of this framework lies in its integrated approach, unlike existing studies that treat these elements separately. It presents a holistic, modular security pipeline with empirically demonstrated effectiveness across multiple performance measures.

The proposed framework achieved a model accuracy of 92%, adversarial robustness of 85%, and a privacy preservation score of 90%, outperforming all referenced works. These results confirm that the framework provides resilience and responsiveness in complex cloud environments while effectively managing model integrity and robustness against adversarial attacks.

Future research will explore real-time AI security monitoring and blockchain-based data integrity mechanisms to enhance reliability, transparency, and compliance in increasingly complex and dynamic cloud AI/ML environments.

REFERENCES

[1] D. Patel *et al.*, "Cloud Platforms for Developing Generative AI Solutions: A Scoping Review of Tools and Services." arXiv, Dec. 08, 2024, https://doi.org/10.48550/arXiv.2412.06044.

[2] A. Habbal, M. K. Ali, and M. A. Abuzaraida, "Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, applications, challenges and future research directions," *Expert Systems with Applications*, vol. 240, Apr. 2024, Art. no. 122442, https://doi.org/10.1016/j.eswa.2023.122442.

[3] A. Saini and R. Sehrawat, "Enhancing Data Security through Machine Learning-based Key Generation and Encryption," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14148–14154, Jun. 2024, https://doi.org/10.48084/etasr.7181.

[4] S. K. Jagatheesaperumal, M. Rahouti, K. Ahmad, A. Al-Fuqaha, and M. Guizani, "The Duo of Artificial Intelligence and Big Data for Industry 4.0: Applications, Techniques, Challenges, and Future Research

Directions," *IEEE Internet of Things Journal*, vol. 9, no. 15, pp. 12861–12885, Aug. 2022, https://doi.org/10.1109/JIOT.2021.3139827.

[5] J. S. Kumar, A. Gupta, S. Tanwar, N. Kumar, and S. Akleylek, "Security enhancement in cellular networks employing D2D friendly jammer for V2V communication," *Cluster Computing*, vol. 26, no. 2, pp. 865–878, Apr. 2023, https://doi.org/10.1007/s10586-022-03551-0.

[6] M. A. M. Farzaan, M. C. Ghanem, A. El-Hajjar, and D. N. Ratnayake, "AI-Enabled System for Efficient and Effective Cyber Incident Detection and Response in Cloud Environments." arXiv, Jan. 12, 2025, https://doi.org/10.48550/arXiv.2404.05602.

[7] J. Robertson, J. M. Fossaceca, and K. W. Bennett, "A Cloud-Based Computing Framework for Artificial Intelligence Innovation in Support of Multidomain Operations," *IEEE Transactions on Engineering Management*, vol. 69, no. 6, pp. 3913–3922, Dec. 2022, https://doi.org/10.1109/TEM.2021.3088382.

[8] V. H. Das Chowdary, A. Shanmukh, T. P. Nikhil, B. S. Kumar, and F. Khan, "DevOps 2.0: Embracing AI/ML, Cloud-Native Development, and a Culture of Continuous Transformation," in *2024 4th International Conference on Pervasive Computing and Social Networking*, Salem, India, 2024, pp. 673–679, https://doi.org/10.1109/ICPCSN62568.2024.00112.

[9] S. Tuli, G. Casale, and N. R. Jennings, "MCDS: AI Augmented Workflow Scheduling in Mobile Edge Cloud Computing Systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 11, pp. 2794–2807, Nov. 2022, https://doi.org/10.1109/TPDS.2021.3135907.

[10] E. Zeydan, S. S. Arslan, and M. Liyanage, "Managing Distributed Machine Learning Lifecycle for Healthcare Data in the Cloud," *IEEE Access*, vol. 12, pp. 115750–115774, 2024, https://doi.org/10.1109/ACCESS.2024.3443520.

[11] P. Abinaya and J. S. Kumar, "Assured and Provable Data Expuncturing in cloud using Ciphertext Policy–Attribute Based Encryption (CP-ABE)," *Cybernetics and Systems*, vol. 55, no. 4, pp. 786–803, May 2024, https://doi.org/10.1080/01969722.2023.2176654.

[12] J. Ejarque *et al.*, "Enabling dynamic and intelligent workflows for HPC, data analytics, and AI convergence," *Future Generation Computer Systems*, vol. 134, pp. 414–429, Sep. 2022, https://doi.org/10.1016/j.future.2022.04.014.

[13] A. Giannopoulos *et al.*, "Supporting Intelligence in Disaggregated Open Radio Access Networks: Architectural Principles, AI/ML Workflow, and Use Cases," *IEEE Access*, vol. 10, pp. 39580–39595, 2022, https://doi.org/10.1109/ACCESS.2022.3166160.

[14] E. e Oliveira, M. Rodrigues, J. P. Pereira, A. M. Lopes, I. I. Mestric, and S. Bjelogrlic, "Unlabeled learning algorithms and operations: overview and future trends in defense sector," *Artificial Intelligence Review*, vol. 57, no. 3, Feb. 2024, Art. no. 66, https://doi.org/10.1007/s10462-023-10692-0.

[15] M. Rahouti, D. Lyons, S. K. Jagatheesaperumal, and K. Xiong, "A Decentralized Cooperative Navigation Approach for Visual Homing Networks," *IT Professional*, vol. 25, no. 6, pp. 71–81, Nov. 2023, https://doi.org/10.1109/MITP.2023.3323865.

[16] "MNIST Dataset." Kaggle. [Online]. Available: https://www.kaggle.com/datasets/hojjatk/mnist-dataset.

[17] W. Cukierski, "CIFAR-10 - Object Recognition in Images." Kaggle, 2013. [Online]. Available: https://kaggle.com/cifar-10.

[18] C. Niloor, R. Agarwal, and P. Mishra, "Using MNIST Dataset for De-Pois Attack and Defence," in *Fourth International Conference on Recent Trends in Communication and Intelligent Systems*, Jaipur, India, 2023, pp. 213–227, https://doi.org/10.1007/978-981-99-5792-7_17.

[19] X. Cao, M. Rahouti, S. K. Jagatheesaperumal, and K. Xiong, "Psychological Information Sharing Using Ethereum Blockchain and Smart Contracts," in *2023 Fifth International Conference on Blockchain Computing and Applications*, Kuwait, Kuwait, 2023, pp. 561–568, https://doi.org/10.1109/BCCA58897.2023.10338936.

[20] M. Uddin, S. Islam, and A. Al-Nemrat, "A Dynamic Access Control Model Using Authorising Workflow and Task-Role-Based Access Control," *IEEE Access*, vol. 7, pp. 166676–166689, 2019, https://doi.org/10.1109/ACCESS.2019.2947377.

[21] T. Zhu, D. Ye, W. Wang, W. Zhou, and P. S. Yu, "More Than Privacy: Applying Differential Privacy in Key Areas of Artificial Intelligence," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 6, pp. 2824–2843, Jun. 2022, https://doi.org/10.1109/TKDE.2020.3014246.

[22] J. Pang and G. Cheung, "Graph Laplacian Regularization for Image Denoising: Analysis in the Continuous Domain," *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1770–1785, Apr. 2017, https://doi.org/10.1109/TIP.2017.2651400.

[23] T. Elgamrani, R. Elgaf, and Y. Chtouki, "Adversarial Attack Defense Techniques: A Study of Defensive Distillation and Adversarial Re-Training on CIFAR-10 and MNIST," in *2024 International Conference on Computer and Applications*, Cairo, Egypt, 2024, pp. 1–4, https://doi.org/10.1109/ICCA62237.2024.10927831.

[24] J. J. Hathaliya, S. Tanwar, and P. Sharma, "Adversarial learning techniques for security and privacy preservation: A comprehensive review," *Security and Privacy*, vol. 5, no. 3, May 2022, Art. no. e209, https://doi.org/10.1002/spy2.209.