IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

PRIVACY-PRESERVING EDGE-AI FRAMEWORK USING TRANSFORMER-BASED FEDERATED LEARNING FOR REAL-TIME INTELLIGENT SURVEILLANCE AND ANOMALY DETECTION

Ganapatiramanan A[1], Dhakshan T[2], Gaurav Dharshan R[3], P Selvamani[4], Dr T Rajasekaran[5]

[1]UG Student, [2]UG Student, [3]UG Student, [4]Assistant Professor, [5]Associate Professor

Department of Computer Science and Engineering, Sri Venkateswara College of Engineering, Sriperumbudur, Tamil Nadu, India

Doi: https://doi.org/10.56975/ijcrt.v14i7.308996

Abstract: The increasing deployment of intelligent surveillance systems has created a need for efficient real-time processing while ensuring data privacy and security. Traditional cloud-centric architectures are often limited by high latency, bandwidth constraints, and risks of sensitive data exposure. This study presents a novel framework that combines edge artificial intelligence, transformer-based deep learning models, and federated learning to enable real-time anomaly detection in surveillance environments. The proposed system processes data at the edge nodes, reducing latency and minimizing the dependency on centralized servers. Federated learning allows collaborative model training without sharing raw data, thereby preserving the privacy of the user. Transformer models enhance detection performance by capturing temporal dependencies in video streams more effectively than conventional methods. The proposed framework demonstrates improved accuracy, reduced response time, and strong privacy guarantees, making it suitable for applications such as smart cities, public safety, and critical infrastructure monitoring.

Keywords **-** Edge AI, Federated Learning, Transformer Models, Privacy Preservation, Surveillance Systems, Anomaly Detection, Deep Learning.

I. INTRODUCTION

The rapid expansion of surveillance technologies in urban and industrial environments has resulted in the generation of large amounts of video data. These systems play a crucial role in ensuring public safety, monitoring activities, and detecting suspicious behaviors. However, conventional surveillance solutions rely heavily on centralized cloud infrastructure, which introduces challenges related to latency, scalability, and data privacy issues. Processing video data in centralized systems can lead to delays that affect real-time decision-making. Additionally, the transmission of sensitive data over networks increases the risk of unauthorized access and privacy violations. These limitations have motivated the adoption of edge computing, in which data processing is performed closer to the source of data generation. Edge artificial intelligence enables local data analysis on devices, such as cameras and edge servers. Although this approach reduces latency, it also introduces challenges in training and updating machine learning models across distributed devices. Federated learning addresses this issue by allowing multiple devices to collaboratively train a shared model without exchanging raw data with each other. Recent advancements in transformer architectures have significantly improved the performance of sequence-based tasks. These models can capture long-range

dependencies in data, making them well-suited for analyzing video streams in surveillance systems. This study proposes a unified framework that integrates edge computing, federated learning, and transformer-based models to create a scalable, efficient, and privacy-preserving surveillance system capable of real-time anomaly detection.

II. RELATED WORK

Early approaches to intelligent surveillance relied on traditional image-processing techniques and simple machine-learning algorithms. With the advancement of deep learning, convolutional neural networks (CNNs) have become widely used for object detection and activity recognition. Recurrent neural networks and long short-term memory models were subsequently introduced to handle temporal data in video sequences. Although these models improve performance, they often require centralized datasets for training, which raises privacy concerns. Additionally, their ability to capture long-term dependencies is limited when addressing complex video patterns. Edge computing has emerged as a solution for reducing latency and improving system responsiveness by processing data locally. However, deploying deep learning models on edge devices presents challenges related to computational resources and model synchronization. Federated learning has gained attention as a decentralized training approach that enables multiple devices to contribute to model training without the need to share sensitive data. This technique has been successfully applied in domains such as healthcare and in mobile applications. Recently, transformer-based models have shown superior performance in handling sequential data owing to their attention mechanisms. Their application in video analysis has opened new possibilities for improving anomaly detection in surveillance systems. Despite these advancements, limited research has combined transformers, federated learning, and edge computing into a single framework for privacy-preserving surveillance. This study addresses this gap by proposing an integrated approach.

III. SYSTEM ARCHITECTURE

The proposed framework comprises three primary layers that work together to achieve efficient and secure surveillance. The first layer is the device layer, which consists of cameras and sensors that are deployed in the environment. These devices capture video data and perform basic preprocessing tasks, such as frame extraction and normalization. The second layer is the edge layer, where core processing occurs. Edge nodes run transformer-based models to analyze video streams and detect anomalies in real-time. The nodes also perform local training using federated learning techniques. The third layer is the coordination layer, which manages the communication between the edge devices. Instead of collecting raw data, this layer aggregates the model updates from different nodes and generates an improved global model. The updated model is then distributed to the edge devices. This architecture ensures that sensitive data remain on local devices while still benefiting from collaborative learning across the network.

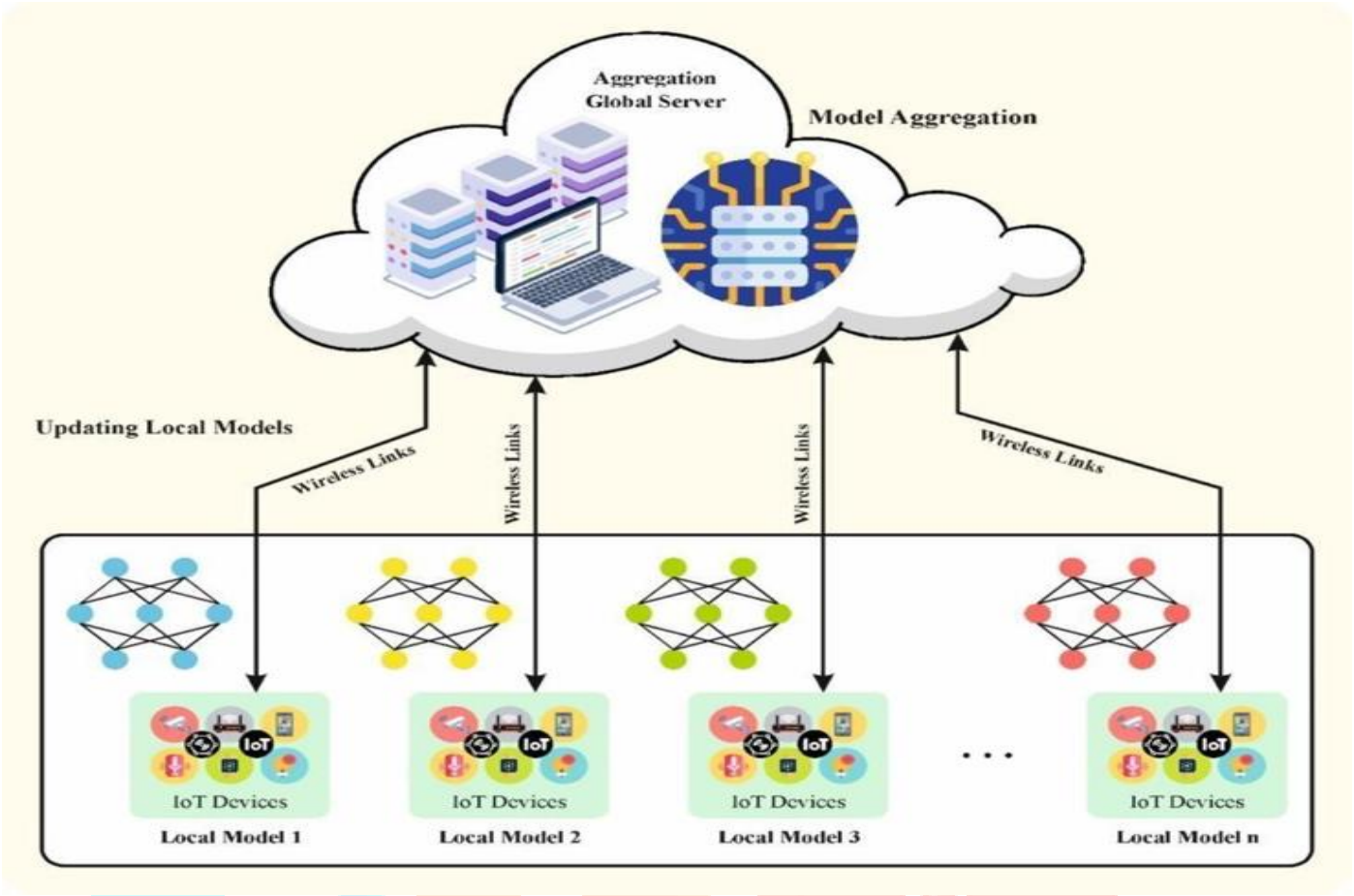

Figure 1: Architecture of Privacy-Preserving Edge-AI Surveillance Framework

The architecture consists of three layers. The IoT device layer includes cameras and sensors that capture real-time data. The edge layer performs local processing using transformer-based models to detect anomalies with low latency. Each edge node trains the model locally and shares only the model parameters with the central server. The cloud coordination layer aggregates updates from multiple edge devices using federated learning and distributes the improved global model back to the nodes. This design ensures efficient processing, scalability, and strong privacy preservation since raw data is not transmitted outside the local devices.

IV. METHODOLOGY

The proposed system comprises multiple stages that ensure efficient processing and accurate anomaly detection. In the preprocessing stage, the video data were divided into frames and cleaned to remove noise. Relevant features were extracted to reduce the computational complexity. The core component of this system is a transformer-based model. Unlike traditional sequential models, transformers use attention mechanisms to analyze the relationships between different parts of the input data. This allows the model to identify unusual patterns in the video sequences more effectively. Federated learning was used to train the model across multiple edge devices. Each device trains the model using its local data and shares only the model parameters with a central coordinator. The coordinator combines these updates to create a global model without accessing the original dataset. Privacy is maintained through secure communication protocols and optional techniques, such as differential privacy. This ensures that information cannot be reconstructed from the shared parameters.

V. IMPLEMENTATION

The performance of the proposed framework was evaluated using several key metrics. Accuracy is measured by comparing the model's ability to correctly identify normal and abnormal events. Transformer-based models demonstrate higher accuracy than traditional approaches because of their ability to capture complex temporal relationships. The latency is significantly reduced because processing occurs at the edge rather than in a centralized cloud. This enables faster response times in critical situations such as cardiac arrest. Privacy is maintained throughout the system because raw data are never transmitted outside the local device. This reduces the risk of data breaches and ensures compliance with privacy regulations. The efficiency is also

improved as the bandwidth usage is minimized. Only model updates are transmitted, which requires significantly less data than raw video streams.

VI. EXPERIMENTAL RESULTS

The performance of the proposed framework was evaluated using several key metrics. Accuracy is measured by comparing the model's ability to correctly identify normal and abnormal events. Transformer-based models demonstrate higher accuracy than traditional approaches because of their ability to capture complex temporal relationships. The latency is significantly reduced because processing occurs at the edge rather than in a centralized cloud. This enables faster response times in critical situations such as cardiac arrest. Privacy is maintained throughout the system because raw data are never transmitted outside the local device. This reduces the risk of data breaches and ensures compliance with privacy regulations. The efficiency is also improved as the bandwidth usage is minimized. Only model updates are transmitted, which requires significantly less data than raw video streams.

Table 1: Comparison between Traditional Surveillance Systems and Proposed Framework

Feature	Traditional Cloud-Based Surveillance	Proposed Edge-AI Framework	Improvement
Data Processing	Centralized(Cloud)	Decentralized (Edge Devices)	Faster processing
Latency	High	Low	Real-time response
Privacy	Low(data shared)	High(no raw data shared)	Secure
Bandwidth Usage	High	Low	Efficient
Real-Time Detection	Limited	Efficient	Better monitoring
Model Training	Centralized	Federated Learning	Distributed
Accuracy	Moderate	High (Transformer-based)	Improved detection
Scalability	Limited	Highly Scalable	Flexible deployment

The effectiveness of the suggested privacy-preserving Edge-AI framework is assessed by contrasting it with conventional cloud-based surveillance systems. Important factors like accuracy, latency, and efficiency are taken into account to evaluate the system's performance. The use of transformer-based models enhances detection abilities, whereas edge computing minimizes response time. Additionally, federated learning improves the system by facilitating collaborative training while safeguarding data privacy.

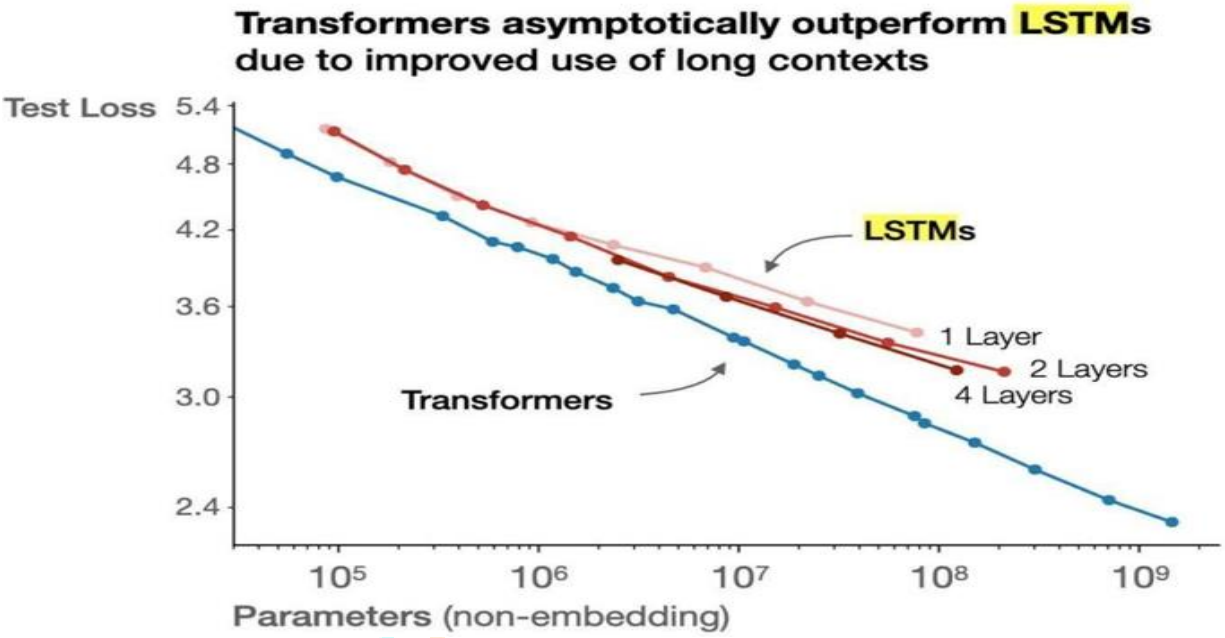

Figure 2: Performance Comparison between Transformer-Based Edge AI and Traditional Models

The graph illustrates the performance comparison between traditional cloud-based systems and the proposed edge AI framework. The proposed system achieves higher accuracy owing to the use of transformer models while maintaining a lower latency through edge processing. This demonstrates the effectiveness of integrating federated learning with edge computing for real-time surveillance.

VII. ADVANTAGES OF THE PROPOSED SYSTEM

The proposed framework offers several advantages over existing approaches. This enables the real-time detection of anomalies, which is essential for surveillance applications. The use of edge computing reduces latency and improves the responsiveness of the system. Privacy is preserved through federated learning, ensuring that sensitive data remain secure. The integration of transformer models enhances the detection accuracy and allows the system to effectively handle complex scenarios. The system is scalable and can be deployed across multiple locations without significant architectural changes. This makes it suitable for large-scale applications, such as smart cities.

VIII. APPLICATIONS

This framework can be applied to various domains that require intelligent monitoring and security. In smart cities, it can be used for traffic management and crime prevention purposes. In industrial environments, it can monitor safety conditions and detect hazardous situations in real time. Public spaces, such as airports, railway stations, and shopping centers, can benefit from real-time anomaly detection. The system can also be used in residential areas for home security and monitoring purposes.

IX. CONCLUSION

This study presents a comprehensive framework for privacy-preserving intelligent surveillance using edge artificial intelligence, federated learning, and transformer-based models. The proposed approach addresses the key challenges associated with traditional surveillance systems, including latency, scalability, and data privacy. By processing data locally and enabling collaborative learning without sharing raw information, the system achieves a balance between performance and security. Experimental analysis showed that the framework provides improved accuracy and efficiency compared with conventional methods. The proposed solution has strong potential for real-world deployment in various surveillance scenarios, particularly in environments where privacy and real-time processing are critical requirements.

X. FUTURE WORK

Future research should focus on improving the efficiency of transformer models for deployment on low-resource edge devices. Additional enhancements may include the integration of explainable artificial intelligence techniques to improve transparency and trust in the system. The use of multimodal data, such as audio and sensor input, can further enhance the anomaly detection capabilities. Real-world testing and

deployment in large-scale environments will also help to validate the effectiveness of the proposed framework.

REFERENCES

. 1. A. Vaswani, Shankar, et al., "Attention Is All You Need," Advances in Neural Information Processing Systems (NeurIPS), 2017.
. 2. H. Brendan McMahan et al., "Communication-Efficient Learning of Deep Networks from Decentralized Data," Proceedings of AISTATS, 2017.
. 3. Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated Machine Learning: Concept and Applications," ACM Transactions on Intelligent Systems and Technology, 2019.
. 4. Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, pp. 436–444, 2015.
. 5. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
. 6. W. Shi, "Edge Computing: Vision and Challenges," IEEE Internet Things Journal, vol. 3, no. 5, pp. 637–646, 2016.
. 7. J. Konečný et al., "Federated Learning: Strategies for Improving Communication Efficiency," arXiv preprint arXiv:1610.05492, 2016.
. 8. A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," International Conference on Learning Representations (ICLR), 2021.
. 9. R. Chalapathy and S. Chawla, "Deep Learning for Anomaly Detection: A Survey," ACM Computing Surveys, 2019.
. 10. T. Li, A. Sahu, A. Talwalkar, and V. Smith, "Federated Learning: Challenges, Methods, and Future Directions," IEEE Signal Processing Magazine, 2020.