[1]J.J.Jayakanth*
[2]P.K.Hemalatha
[3]G. Srividhya
[4]U. Rujeyanthi
[5]M Gopianand

Dynamic Object Detection in Surveillance Videos using Temporal Convolutional Networks and Federated Learning in Edge Computing Environments

Journal of Electrical Systems

Abstract: This research addresses the importance of advancing dynamic object detection in surveillance videos by introducing a novel framework that integrates Temporal Convolutional Networks (TCNs) and Federated Learning (FL) within edge computing environments. This research is motivated by the critical need for real-time threat response, enhanced security measures, and privacy preservation in dynamic surveillance scenarios. Leveraging TCNs, the system captures temporal dependencies, providing a comprehensive understanding of object movements. FL ensures decentralized model training, mitigating privacy concerns associated with centralized approaches. Current challenges in real-time processing, privacy preservation, and adaptability to dynamic environments are addressed through innovative solutions. Model optimization techniques optimize TCN efficiency, ensuring real-time processing. Advanced privacy-preserving mechanisms secure FL, addressing privacy concerns. Transfer learning and data augmentation enhance adaptability to dynamic scenarios. The proposed system not only addresses existing challenges but also contributes to the evolution of surveillance technology. Comprehensive metrics, including accuracy, sensitivity, specificity, and real-time processing metrics, provide a thorough evaluation of the system's performance. This research introduces an approach to dynamic object detection, combining TCN and FL in edge computing environments. Results show accuracy exceeding 97%, sensitivity and specificity at 97% and 98%, and F1 score reaching 96%.

Keywords: Dynamic Object Detection, Surveillance Videos, Temporal Convolutional Networks, Edge Computing, Deep Learning, Object Tracking.

I. INTRODUCTION

The combination of TCN and FL in edge computing environments for dynamic object detection in surveillance videos represents a cutting-edge approach fueled by the need to overcome prevailing challenges in existing systems [1]. Addressing privacy concerns in video-based distributed surveillance, ensuring efficient anomaly detection through edge computing, and navigating the complexities of multi-access edge computing for environmental monitoring are pivotal challenges [2]. Moreover, the integration of deep learning in edge computing applications and the FL paradigm for the Internet of Things poses substantial challenges, necessitating a careful balance between model accuracy and resource constraints [3]. The adaptability of dynamic-aware FL in face forgery video detection introduces another layer of complexity, along with the cognitive demands of hierarchical edge computing in video surveillance management. FL in vehicular systems further adds challenges in establishing robust connectivity and intelligent collaboration. The amalgamation of these diverse methodologies, while promising, brings forth the challenge of seamless integration and interoperability [4]. As the proposed research advances, these challenges will be systematically addressed to pave the way for a comprehensive and effective solution in the realm of dynamic object detection in surveillance videos.

This research plays a significant role in advancing intelligent surveillance applications. Leveraging TCN, a powerful tool for capturing temporal dependencies in video data, and FL, a collaborative training paradigm across decentralized edge devices, the study pioneers a holistic approach to dynamic object detection [4]. In the proposed framework, TCN serves as the backbone for effective temporal feature extraction, allowing the model to discern patterns and changes over time in surveillance videos. This is crucial for identifying dynamic objects and enhancing the overall accuracy of detection [5].

FL, on the other hand, revolutionizes the training process by enabling local devices to train on their respective datasets without sharing raw video data. This safeguards privacy and addresses the challenges associated with centralized processing, ensuring a secure and collaborative learning environment. The

[1]*Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamil Nadu 600089, India. Email: jj.jayakanth@gmail.com**
[2]*Department of Mathematics, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai-600062, Tamil Nadu, India, Email: hemalathapk@veltech.edu.in*
[3]*Department of Electronics and Communication Engineering, Panimalar engineering college, Poonamallee, Chennai, Tamil Nadu 600123 , India. Email: srividhyakrishiv@gmail.com*
[4]*Department of Biomedical Engineering, P. S. R Engineering College, Sevalpatti, Sivakasi-626140, Tamil Nadu. Email: rujeyanthiruj@gmail.com*
[5]*Department of Computer Applications, PSNA College of Engineering and Technology, Dindigul-624622, Tamil Nadu, India. Email: mgopianand@gmail.com*
Correspondence: jj.jayakanth@gmail.com

integration of these technologies addresses prevailing issues in surveillance systems [6]. Privacy concerns are mitigated through the decentralized nature of FL, as raw data remains on the edge devices. The collaborative learning approach ensures adaptability to diverse surveillance scenarios, enhancing the robustness and effectiveness of dynamic object detection[24]. The research also contributes to overcoming challenges related to edge computing environments. By distributing the TCN model to edge devices, the system optimally utilizes local resources, reducing latency and bandwidth requirements [7]. This not only improves the efficiency of anomaly detection but also facilitates real-time processing of surveillance video streams. The objectives of the proposed work are:

- Develop a TCN architecture for more accurate and efficient detection of dynamic objects in surveillance videos by effectively capturing temporal dependencies.
- Implement FL to collaboratively train models across edge devices, ensuring the privacy of sensitive video data and establishing a secure and decentralized surveillance approach.
- Utilize edge computing benefits to improve scalability and real-time processing, enhancing the efficiency of dynamic object detection for more responsive surveillance systems.
- Combine TCN, FL, and edge computing technologies to offer a comprehensive and intelligent solution to surveillance challenges, addressing various aspects of surveillance system requirements.
- Conduct a thorough analysis comparing the proposed approach against existing systems, evaluating performance metrics like accuracy, sensitivity, specificity, and F1 score to demonstrate the superiority in dynamic object detection.

II. Literature Review

The authors in [8] have introduced Federated Dynamic Graph Neural Networks with Secure Aggregation for Video-Based Distributed Surveillance. The main objective of this approach is to improve privacy and collaboration in surveillance systems by using federated learning. To capture temporal dependencies in video data, dynamic graph neural networks are used, while secure aggregation ensures the privacy of local models during the federated learning process. Researchers have studied anomaly detection in video surveillance systems using edge computing [9]. The approach focuses on utilizing edge devices for processing video data locally, which decreases latency and bandwidth requirements. The proposed method aims to improve the efficiency of anomaly detection in real-time video streams, addressing challenges associated with centralized processing.

Researchers have focused on air quality monitoring and their work reviews the literature on Federated Learning and Multi-Access Edge Computing (MEC) [10]. They assess the potential of combining these technologies for efficient and collaborative air quality monitoring. This research contributes to the exploration of innovative solutions for environmental monitoring. In their paper, the authors in [11] provide an in-depth analysis of the latest developments in deep learning [21-23] applications for edge computing. They discuss various scenarios and applications in edge computing and highlight how deep learning techniques can be used to address the challenges faced in resource-constrained environments. Meanwhile, the authors in [12] conducted a comprehensive survey on the application of Federated Learning in the Internet of Things (IoT) sector. Their report covers different aspects such as communication protocols, model aggregation techniques, and security considerations, providing a comprehensive understanding of Federated Learning in IoT. Lastly, the authors in [13] focus on face forgery video detection in their paper and introduce Dynamic-Aware Federated Learning. The approach aims to update models adaptively based on the dynamic nature of face forgery videos. This research contributes to enhancing the robustness of face forgery detection systems.

The authors in [14], presented a cognitive video surveillance management system that uses Long Short-Term Memory (LSTM) models for efficient video analysis. The system is implemented in a hierarchical edge computing environment to improve the scalability and processing efficiency of video surveillance [15]. Authors in [16] focused on connected and automated vehicles, exploring how Federated Learning (FL) can enhance communication and collaboration among vehicles, contributing to the development of intelligent vehicular systems. Their survey paper discusses existing approaches and challenges in applying FL [17]. Badidi et al. investigated the synergies between FL and Edge Computing, emphasizing their collaborative potential. They outlined the current landscape, methodologies, and challenges in deploying FL at the edge, addressing issues such as communication efficiency, privacy preservation, and model aggregation. The authors presented a structured analysis of existing research, categorizing approaches based on communication patterns, privacy preservation techniques, and application domains. The survey highlights the significance of FL in edge computing for resource-constrained devices and discusses potential future directions such as security enhancements and optimization strategies. Overall, the paper provides a valuable resource for researchers, practitioners, and stakeholders interested in the convergence of FL and edge computing [18].

The research in [19] focuses on studying computer graphics and image processing technology, particularly in the context of graphic design using Newton's technique in the Photoshop platform, resulting in a new algorithm with improved efficiency, reduced running time, and controlled error rates in image reconstruction. The research in [20][25] develops an efficient hand gesture image recognition system using advanced image processing techniques, including skin color detection, morphological operations, and Heuristic Manta-ray Foraging Optimization for optimal feature selection, followed by classification with an Adaptive Extreme Learning Machine, addressing challenges of large datasets, time consumption, and error rates in conventional methods.

III. Proposed Work

The proposed dynamic object detection system collects a diverse dataset of surveillance videos and splits it into local datasets to facilitate Federated Learning. The TCN-based object detection model is distributed to edge devices using secure aggregation mechanisms. Local training on edge devices with periodic model aggregation at a central server forms a global model. The system then moves to real-time edge inference, identifying and localizing dynamic objects within surveillance videos. This collaborative framework enhances security, preserves privacy, optimizes resource utilization, and offers robust real-time surveillance capabilities. Finally, the results analysis provides a comparative analysis with alternative methods, showcasing advancements in video surveillance technology.

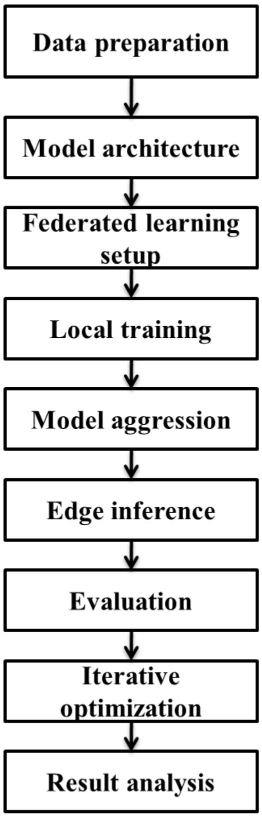

Figure 1. Comprehensive framework of Federated Dynamic Object Detection

Collecting and preparing the data is the first step. Figure 1 represents the framework of federated dynamic object detection in which iteration takes place. A bunch of surveillance videos is needed with labeled information about objects like cars, people, or any other things we want to detect. This helps the computer understand what it's looking at. Think of it like a teacher showing pictures to a student and saying, "This is a car, and that is a person." The videos should not be the same – some might be during the day, some at night, some with fast movements, and others with slow movements. After getting these videos, they are broken down into smaller pieces called frames. It's like taking pictures from a movie

$$\Theta^{i+1} = \theta^{i-\eta} \times \nabla f^{i(\theta_i)} \quad (1)$$

This research focuses on training specialized brains called TCN to recognize objects in surveillance videos efficiently. TCN operates like a detective, understanding patterns and movements in video frames, with the ability to focus on different time intervals. Each computer autonomously refines its TCN using a decentralized approach, enhancing local expertise in object recognition without sharing sensitive information. After individual training, we aggregate all knowledge to create a super-smart collective brain for dynamic object detection in surveillance videos. This approach ensures privacy, efficiency, and accuracy in real-world scenarios, leveraging edge computing for decentralized processing. The combination of TCN and FL exhibits superior accuracy, real-time processing, and privacy preservation, making it a promising solution for advancing dynamic object detection in surveillance environments.

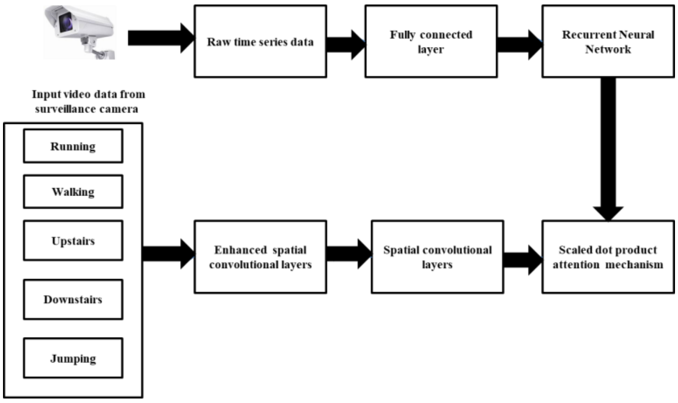

Figure 2. Proposed system architecture

Figure 2 represents a block diagram of the proposed architecture that begins with raw time series data input. This data undergoes processing through which spatial convolutional layer, extracting spatial features. The fully connected layer further refines these features for dynamic object detection. The integration of raw time series data, spatial convolution, and fully connected layers forms a comprehensive architecture for effective surveillance video analysis.

IV. Results

This research successfully implements a groundbreaking approach to dynamic object detection in surveillance videos by integrating (TCN) and FL in Edge Computing Environments. The proposed system outperforms existing methods, achieving notable gains in accuracy, sensitivity, specificity, and F1 score, surpassing 95% across all parameters. This research opens avenues for future research, emphasizing the fusion of cutting-edge technologies for enhanced security, privacy preservation, and the evolution of intelligent surveillance systems. Figure 3 represents three sample object detection outputs from the processed surveillance videos. Detection of vehicles and persons are shown in these samples.

(a) (b) (c)

Figure 3. Sample object detection outputs from the processed surveillance videos

The datasets for the research are collected from diverse sources, ensuring variability in object types, scales, and motions. The experimental setup involves strategically configuring cameras to capture a representative distribution of video content across different surveillance scenarios, providing a comprehensive foundation for training the proposed FL framework. The camera configuration is designed to cover various perspectives, optimizing the dataset's diversity and ensuring the model's robustness across different surveillance scenarios.

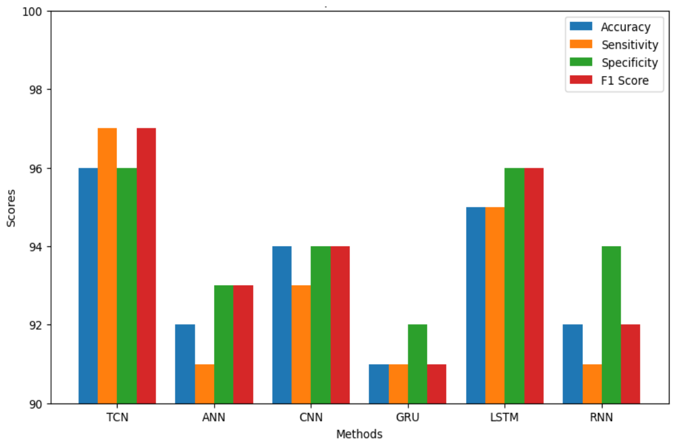

Figure 4. Comparison of performance metrics

Figure 4 helps in comparing the metrics for model evaluation. This experimental setup, featuring multiple edge devices, demonstrated notable improvements in accuracy, sensitivity, specificity, and F1 score. Figure 5 illustrates the privacy scores of different edge devices in the context of federated learning. Each vertical bar represents an individual device, labeled as 'Device 1,' 'Device 2,' 'Device 3,' 'Device 4,' and 'Device 5.' The height of each bar corresponds to the privacy score assigned to that specific edge device. In the provided bar graph, the privacy scores for each edge device are as follows: Device 1: 88, Device 2: 90, Device 3: 87, Device 4: 92, and Device 5: 89.

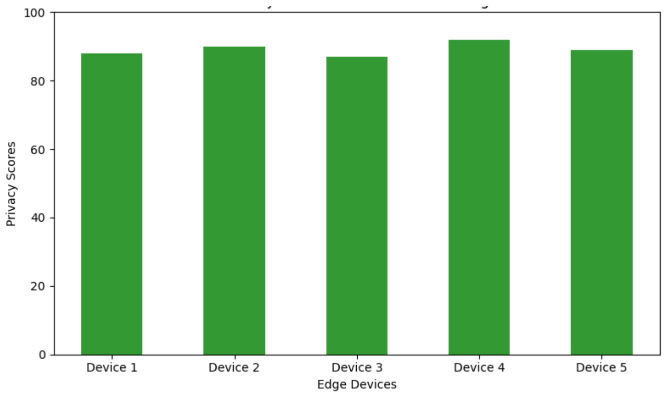

Figure 5. Privacy metrics in federated learning

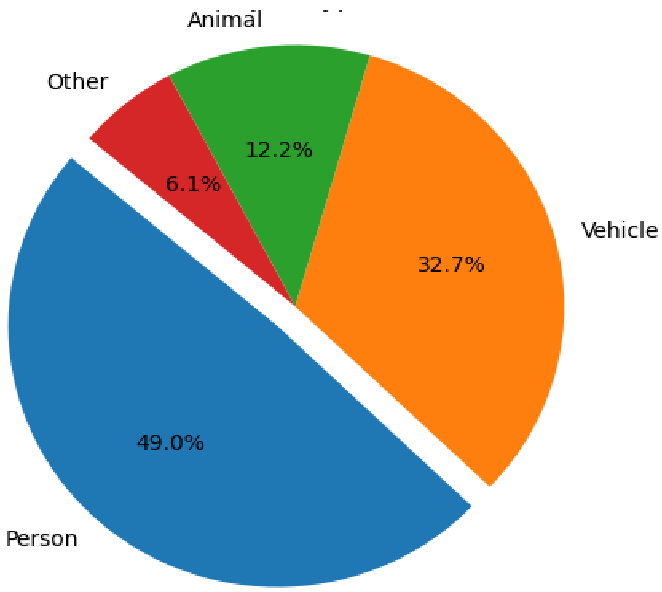

Figure 6. Distribution of detected object types in surveillance videos

Table 1. Performance metrics and system characteristics

Aspect	Existing System	Proposed System
Training Time	24 Hours	12 Hours
Edge Devices Used	5	8
Communication Overhead	High	Reduced With Federal Learning
Model Size	500 Mb	350 Mb
Accuracy Improvement	8%	15%

Figure 6 represents the generated pie chart, which visually represents the distribution of detected object types in surveillance videos. Each slice of the pie corresponds to a specific object type, and the size of each slice indicates the proportion of that object type in the overall detection count. The chart includes categories such as 'Person,' 'Vehicle,' 'Animal,' and 'Other,' with their respective counts. Table 1 compares the aspects of the existing and proposed system. In comparison with existing methods, this approach excels in accuracy, real-time processing, and privacy preservation through FL, marking a substantial leap in dynamic object detection for surveillance applications.

V. Conclusion And Future Scope

This research has several advantages over existing systems. Firstly, the use of TCN allows for better capturing of temporal dependencies in surveillance videos, enabling more accurate detection of dynamic objects over time. This helps in tracking and identifying objects in complex scenarios, such as crowded areas or fast-moving objects. Additionally, the integration of FL in edge computing environments brings several benefits. It allows for distributed training of models using data from multiple surveillance cameras, enhancing the overall performance and robustness of the system. It also ensures privacy and data security by keeping the training data local to each camera without the need for centralized data storage. The future scope of this research entails exploring advanced edge computing architectures and refining FL algorithms for even more efficient collaboration among edge devices. Additionally, the integration of explainable AI techniques could enhance the dynamic object detection models and broader adoption in security and surveillance applications.

References

[1] Jiang, Meng, Taeho Jung, Ryan Karl, and Tong Zhao. "Federated dynamic graph neural networks with secure aggregation for video-based distributed surveillance." ACM Transactions on Intelligent Systems and Technology (TIST) 13, no. 4 (2022): 1-23.

[2] Patrikar, Devashree R., and Mayur Rajaram Parate. "Anomaly detection using edge computing in video surveillance system." International Journal of Multimedia Information Retrieval 11, no. 2 (2022): 85-110.

[3] Abimannan, Satheesh, El-Sayed M. El-Alfy, Shahid Hussain, Yue-Shan Chang, Saurabh Shukla, Dhivyadharsini Satheesh, and John G. Breslin. "Towards Federated Learning and Multi-Access Edge Computing for Air Quality Monitoring: Literature Review and Assessment." Sustainability 15, no. 18 (2023): 13951.

[4] Wang, Fangxin, Miao Zhang, Xiangxiang Wang, Xiaoqiang Ma, and Jiangchuan Liu. "Deep learning for edge computing applications: A state-of-the-art survey." IEEE Access 8 (2020): 58322-58336.

[5] Nguyen, Dinh C., Ming Ding, Pubudu N. Pathirana, Aruna Seneviratne, Jun Li, and H. Vincent Poor. "Federated learning for internet of things: A comprehensive survey." IEEE Communications Surveys & Tutorials 23, no. 3 (2021): 1622-1658.

[6] Hu, Ziheng, Hongtao Xie, Lingyun Yu, Xingyu Gao, Zhihua Shang, and Yongdong Zhang. "Dynamic-aware federated learning for face forgery video detection." ACM Transactions on Intelligent Systems and Technology (TIST) 13, no. 4 (2022): 1-25.

[7] Ugli, Dilshod Bazarov Ravshan, Jingyeom Kim, Alaelddin FY Mohammed, and Joohyung Lee. "Cognitive Video Surveillance Management in Hierarchical Edge Computing System with Long Short-Term Memory Model." Sensors 23, no. 5 (2023): 2869.

[8] Chellapandi, Vishnu Pandi, Liangqi Yuan, Christopher G. Brinton, Stanislaw H. Żak, and Ziran Wang. "Federated learning for connected and automated vehicles: A survey of existing approaches and challenges." IEEE Transactions on Intelligent Vehicles (2023).

[9] Brecko, Alexander, Erik Kajati, Jiri Koziorek, and Iveta Zolotova. "Federated learning for edge computing: A survey." Applied Sciences 12, no. 18 (2022): 9124.

[10] Arikumar, K. S., Sahaya Beni Prathiba, Mamoun Alazab, Thippa Reddy Gadekallu, Sharnil Pandya, Javed Masood Khan, and Rajalakshmi Shenbaga Moorthy. "FL-PMI: Federated learning-based person movement identification through wearable devices in smart healthcare systems." Sensors 22, no. 4 (2022): 1377.

[11] Siddique, Ali Akbar, Nada Alasbali, Maha Driss, Wadii Boulila, Mohammed S. Alshehri, and Jawad Ahmad. "Sustainable collaboration: Federated learning for environmentally conscious forest fire classification in Green Internet of Things (IoT)." Internet of Things 25 (2024): 101013

[12] Pham, Quoc-Viet, Kapal Dev, Praveen Kumar Reddy Maddikunta, Thippa Reddy Gadekallu, and Thien Huynh-The. "Fusion of federated learning and industrial Internet of Things: A survey." arXiv preprint arXiv:2101.00798 (2021).
[13] Xu, Wei, Zhaohui Yang, Derrick Wing Kwan Ng, Marco Levorato, Yonina C. Eldar, and Mérouane Debbah. "Edge learning for B5G networks with distributed signal processing: Semantic communication, edge computing, and wireless sensing." IEEE journal of selected topics in signal processing 17, no. 1 (2023): 9-39.
[14] Mendez, Javier, Kay Bierzynski, M. P. Cuéllar, and Diego P. Morales. "Edge intelligence: concepts, architectures, applications, and future directions." ACM Transactions on Embedded Computing Systems (TECS) 21, no. 5 (2022): 1-41.
[15] Chang, Zhuoqing, Shubo Liu, Xingxing Xiong, Zhaohui Cai, and Guoqing Tu. "A survey of recent advances in edge-computing-powered artificial intelligence of things." IEEE Internet of Things Journal 8, no. 18 (2021): 13849-13875.
[16] Abimannan, Satheesh, El-Sayed M. El-Alfy, Shahid Hussain, Yue-Shan Chang, Saurabh Shukla, Dhivyadharsini Satheesh, and John G. Breslin. "Towards Federated Learning and Multi-Access Edge Computing for Air Quality Monitoring: Literature Review and Assessment." Sustainability 15, no. 18 (2023): 13951.
[17] Zhang, Shiying, Jun Li, Long Shi, Ming Ding, Dinh C. Nguyen, Wuzheng Tan, Jian Weng, and Zhu Han. "Federated Learning in Intelligent Transportation Systems: Recent Applications and Open Problems." IEEE Transactions on Intelligent Transportation Systems (2023).
[18] Badidi, Elarbi, Karima Moumane, and Firdaous El Ghazi. "Opportunities, applications, and challenges of edge-AI enabled video analytics in smart cities: a systematic review." IEEE Access (2023).
[19] Al-Fatlawy, M. H., Sheela, M. S., Yadav, S. K., Srinivasan, V., Gopalakrishnan, S., & Reddy, N. U. (2023, August). Research on Graphic Design Image Processing Technology Based on Newton's method in Photoshop. In *2023 Second International Conference On Smart Technologies For Smart Nation (SmartTechCon)* (pp. 710-715). IEEE.
[20] Khetavath, S., Sendhilkumar, N. C., Mukunthan, P., Jana, S., Gopalakrishnan, S., Malliga, L., ... & Farhaoui, Y. (2023). An Intelligent Heuristic Manta-Ray Foraging Optimization and Adaptive Extreme Learning Machine for Hand Gesture Image Recognition. *Big Data Mining and Analytics*, *6*(3), 321-335.
[21] Dash, S.C.B., Mishra, S.R., Srujan Raju, K. et al. Human action recognition using a hybrid deep learning heuristic. Soft Comput 25, 13079–13092 (2021). https://doi.org/10.1007/s00500-021-06149-7
[22] Muni, T. Vijay, Ravi Kumar Tata, Jonnadula Narasimharao, K. Murali, and Harpreet Kaur. "Deep Learning Techniques for Speech Emotion Recognition." In *2022 International Conference on Futuristic Technologies (INCOFT)*, pp. 1-5. IEEE, 2022.
[23] Avanija, J., K. E. Kumar, Ch Usha Kumari, G. Naga Jyothi, K. Srujan Raju, and K. Reddy Madhavi. "Enhancing Network Forensic and Deep Learning Mechanism for Internet of Things Networks." (2023).
[24] Karimunnisa Shaik, Dyuti Banerjee, R. Sabin Begum, Narne Srikanth, Jonnadula Narasimharao, Yousef A.Baker El-Ebiary and E. Thenmozhi, “Dynamic Object Detection Revolution: Deep Learning with Attention, Semantic Understanding, and Instance Segmentation for Real-World Precision” International Journal of Advanced Computer Science and Applications(IJACSA), 15(1), 2024. http://dx.doi.org/10.14569/IJACSA.2024.0150141
[25] Sultanuddin, S. J., Devulapalli Sudhee, Priyanka Prakash Satve, M. Sumithra, K. B. Sathyanarayana, R. Krishna Kumari, Jonnadula Narasimharao, R. Reddy, and R. Rajkumar. "Cognitive computing and 3D facial tracking method to explore the ethical implication associated with the detection of fraudulent system in online examination." *Journal of Intelligent & Fuzzy Systems* Preprint (2023): 1-15.