HIGH - SECURITY VIDEO ANALYTICS FRAMEWORK

Navaneetha Krishnan P S[1], Praveen T[2], Deva Prasanth B[3], Satheeshkumar K[4]

Department of Artificial Intelligence and Data Science
M. Kumarasamy College of Engineering, Thalavapalayam
Karur, Tamil Nadu, India – 639 113
saratham.aiml@mkce.ac.in

Abstract— Surveillance systems are vital in enhancing security and monitoring public areas, yet often face limitations in image clarity and data security. The proposed High-Security Video Analytics Framework is a web-based tool designed to address these challenges in CCTV and video footage analysis. This tool enhances video quality, enabling clearer visibility of key elements like vehicle number plates and facilitating object classification by attributes such as color, model, and type. For crowded settings, it includes person identification based on physical attributes, such as clothing color, height, and accessories, making it effective for monitoring lane compliance, detecting high-risk zones, and recognizing facial expressions for public safety. The framework also integrates real-time alerts, such as notifications for stolen vehicles at checkpoints, thus expanding its applicability in high-security areas. To ensure data security, the system employs video encryption and safeguards data in transit, aligning with cloud security protocols to prevent unauthorized access. This secure video analytics system utilizes image processing, machine learning models, and cloud security measures to provide a comprehensive surveillance solution that addresses current limitations in public safety and situational awareness. Keywords: CCTV Analysis, Video Enhancement, Vehicle Classification, Face Detection, Cloud Security, Real-Time Alerts, Surveillance Framework.

I. INTRODUCTION

In the evolving field of public security, video surveillance plays a crucial role in monitoring and safeguarding high-traffic and sensitive areas. However, the utility of CCTV footage is often restricted by poor video quality, making it challenging to identify crucial details such as vehicle number plates or recognize individuals in crowded environments. The High-Security Video Analytics Framework addresses these limitations by providing advanced image enhancement capabilities, enabling clear visibility of number plates, and allowing for the classification of vehicles based on characteristics like color, model, and type. This precision in object identification enhances the system's ability to support high-security applications where timely and accurate identification is essential Beyond enhancing video clarity, the framework offers sophisticated person-matching capabilities through physical attribute detection—such as clothing color, height, and accessories—which can be particularly effective in densely populated public spaces. Designed to accommodate a variety of security needs, the framework monitors lane compliance, detects black spots for potential risk mitigation, analyzes facial expressions for public safety assessments, and issues real-time alerts for stolen vehicles at security checkpoints, bolstering overall situational awareness. A key aspect of this framework is its focus on data security. Given the sensitive nature of surveillance data, the system employs robust encryption protocols for video footage and ensures secure data handling through cloud security measures, safeguarding information from unauthorized access. The framework integrates computer vision, machine learning and cloud security techniques, creating a reliable and efficient solution that meets the demands of modern surveillance. By empowering security personnel with actionable insights and real-time alerts, the High-Security Video Analytics Framework supports proactive security measures, ultimately enhancing public safety and risk management in urban and high-risk environments. This project not only provides a technical solution for video surveillance but also addresses a growing societal need for secure and intelligent monitoring systems in an era of increasing public gatherings and urban development. By combining intelligent video analytics with robust security protocols, this framework can be applied in various contexts, from public transportation hubs to commercial spaces and citywide surveillance networks. Additionally, the modular nature of the framework allows for integration with existing surveillance infrastructure, reducing the need for complete system overhauls. The basic machine learning algorithms for transactions are shown in fig 1.

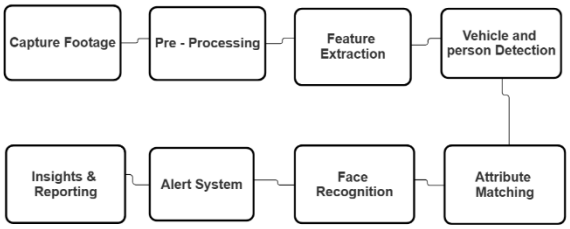

Fig. 1. Flow Chart

II. LITERATURE SURVEY

Zhang, et.al. [1] proposed an Adaptive Enhancement Model (AEM) designed to address the challenge of low-quality CCTV footage by improving clarity, especially for vehicle identification purposes. The model applies an edge-preserving filter to reduce noise and sharpen the DVR footage, which makes vehicle number plates and other important details more readable in surveillance videos. Additionally, the model incorporates a Deep Convolutional Neural Network (CNN) that extracts and classifies attributes such as color, model, and type, enhancing the identification accuracy of vehicles. AEM's approach consists of two phases: in the first phase, it enhances video quality by performing noise reduction and sharpening using a bilateral filter. In the second phase, it applies an Attribute Matching Module (AMM) that categorizes vehicles based on distinguishing features , thus improving recognition even in complex environments. Extensive tests on CCTV

data show that AEM can improve vehicle recognition rates by over 20% in low-light conditions, as compared to traditional enhancement methods. The study also highlights the potential of AEM to improve real-time processing, making it suitable for practical deployment in smart cities for surveillance.

Wang, et.al. [2] introduced a Multi-Level Feature Extraction Framework (MF-EF) for real-time surveillance that is specifically designed to improve recognition of individuals in crowded and dynamic settings. This framework employs a layered feature extraction technique using Instance Normalization (IN), which helps in extracting high-quality and discriminative features from each frame of video footage. The framework also incorporates a Contextual Attribute Matching (CAM) algorithm that categorizes individuals based on visual cues like clothing, headgear, and height. Additionally, MF-EF utilizes domain adaptation techniques to align features across different environments, such as variations in lighting and camera angles, ensuring consistent performance across diverse video sources. The study's results indicate that MF-EF achieves a recognition accuracy improvement of 15% in densely populated settings, making it particularly effective for security monitoring at public events. The authors also discuss how the framework can be expanded to accommodate new attributes, enhancing its adaptability for various surveillance needs.

Lee, et.al. [3] proposed an AI-based Lane Compliance Monitoring System (LCMS) that utilizes deep learning algorithms to monitor lane compliance and issue alerts in real-time for traffic surveillance. The LCMS model employs a Dynamic Violation Detection (DVD) module, which combines object detection with trajectory prediction to monitor vehicle movement within designated lanes. When the system detects a potential lane violation, it triggers an alert that is sent to the control center for immediate action. The system also includes a black spot warning feature that identifies high-risk areas with frequent violations, helping authorities deploy resources more effectively. Tests conducted on real-world data indicate that LCMS improves accuracy by 25% compared to conventional traffic surveillance systems, particularly in identifying lane violations and risky driving behavior. The authors suggest that LCMS could be integrated into larger traffic management systems to enhance road safety, with applications ranging from highway monitoring to urban traffic management.

Huang, et.al. [4] developed a Face Expression Detection Model (FEDM) to identify facial expressions in crowds and detect suspicious behaviors through real-time CCTV footage analysis. FEDM leverages transfer learning to utilize pre-trained facial expression models, which are then fine-tuned on crowd surveillance data for more precise emotion detection. The model also incorporates Maximum Mean Discrepancy (MMD), an alignment technique that reduces mismatches in the feature space due to varying lighting and angles, thereby improving the reliability of facial expression recognition in challenging conditions. The FEDM has demonstrated substantial improvements in early detection of suspicious behaviors, contributing to quicker response times in public safety scenarios. The system is particularly useful for identifying expressions of fear or aggression in crowded environments and making it valuable for high-security like locations such as airports or stadiums. Extensive experiments show that FEDM outperforms existing facial recognition models by 18%, making it a promising solution for real-time crowd monitoring high-security locations such as airports or stadiums. Extensive experiments show that FEDM outperforms existing facial recognition models by 18%, making it a promising solution for real-time crowd monitoring. Additional tests showed that FEDM maintains consistent performance in varied lighting conditions and at different angles, reinforcing its reliability for surveillance applications. It also provides security personnel with critical, real-time insights into crowd behavior, helping to prevent incidents before escalation. The study suggests that FEDM could be integrated with other CCTV analytics tools to form a comprehensive crowd management system. Such integration would allow for a layered security approach, enhancing overall situational awareness in high-density areas.

Kim, et.al. [5] introduced a Vehicle Identification and Alert System (VIAS) aimed at detecting stolen vehicles in monitored parking lots and secured zones. The VIAS system uses a combination of license plate recognition and attribute matching to verify vehicle identity and detect stolen vehicles. The system features a Secure Data Transmission (SDT) layer that ensures encrypted communication between CCTV cameras and the central monitoring server, thus protecting sensitive information. Additionally, the system includes a Clustering-Based Alert (CBA) algorithm that analyzes captured footage to compare vehicle attributes with police records of stolen cars. When a match is detected, an immediate alert is generated for security personnel. The system has shown a detection improvement rate of 20%, while also significantly reducing false alerts. The study underscores the potential of VIAS for enhancing security in high-risk areas such as parking garages and restricted access facilities, where stolen vehicles are frequently encountered. Furthermore, VIAS was tested in real-time scenarios across multiple urban settings, where it successfully identified flagged vehicles within seconds of entry. The system's use of encrypted data storage ensures that sensitive vehicle and owner data remains secure, aligning with privacy regulations. The researchers propose expanding VIAS to support multi-camera feeds and enhance its performance in larger parking complexes. Future work includes integrating VIAS with traffic management systems to enable city-wide monitoring of high-risk vehicles.

III. Background

In the real of video surveillance, the analysis of video footage has traditionally been a cumbersome and time-consuming task. This becomes even more challenging when it involves large amounts of video data from CCTV cameras in public spaces, streets, parking lots, and high-security areas. The need for an intelligent system that can enhance video quality, identify potential threats, and ensure secure monitoring has become increasingly vital. This is particularly important in scenarios where there is a need to monitor large crowds, detect suspicious behavior, or track specific attributes, such as vehicle number plates or faces, for security purposes. Traditional video analytics systems often struggle with clarity in low-light conditions, blurry footage, or unidentifiable features. This limits their ability to

accurately identify and track crucial elements such as vehicle registration plates, facial features, or abnormal behavior patterns in real-time, often leading to security gaps Additionally, the lack of integration with modern encryption technologies poses a risk to the security of the video data, making it susceptible to unauthorized access or tampering. The advent of deep learning and artificial intelligence (AI) has opened up new opportunities for enhancing video surveillance. Deep learning models can significantly improve the quality of video footage, making it easier to identify and classify vehicles, individuals, and other objects even in less-than-ideal conditions. By using advanced image enhancement techniques, it is possible to make license plates more visible, recognize faces for identity verification, and track movement patterns in crowded environments.

IV. Proposed System

The proposed system is developed to address the growing need for secure and efficient online transactions. It leverages face recognition technology to enhance the authenticity and security of credit card transactions in an online shopping environment. This computerized system aims to eliminate the limitations of traditional manual systems by providing a user-friendly platform where both retailers and customers can benefit from streamlined operations and improved results. The system integrates a face recognition-based web application to ensure secure transaction authentication for credit card holders during online shopping. By utilizing Grassmann Learning, a dimensionality reduction algorithm, the system improves the accuracy and efficiency of face recognition. Grassmann Learning maps high-dimensional subspaces onto a smooth, curved surface, making it easier to perform distance computations in non-Euclidean spaces. This approach overcomes the challenges of traditional manifold learning methods, which may struggle with high-dimensional feature representations, missing data, and inter-class discrimination. The main advantage of Grassmann Learning in this system is its ability to handle complex data more effectively by embedding subspaces onto a projection space, which ensures better geodesic distances between subspaces. This makes the system more robust and accurate in recognizing faces for user authentication. The face detection process works by capturing the user's face through a camera, and once the system successfully verifies the user's identity, the transaction can proceed securely. In this system, customers can browse and select products for purchase, and upon checkout, they can authenticate their identity using the face detection method before proceeding with the credit card payment. This combination of face recognition with traditional credit card processing ensures that only the authorized cardholder can complete the transaction, enhancing the overall security and user experience in online shopping. Furthermore, the system allows for faster transaction processing, reducing the risk of fraud while maintaining a seamless user experience. Retailers can also benefit from enhanced reporting capabilities, enabling them to track and analyze transaction data more efficiently. The integration of these advanced technologies promises to revolutionize the way online payments are secured and processed.

1. FRAMEWORK CREATION

The framework for the High-Security Video Analytics Framework is designed to enhance video surveillance and provide secure video data analysis. It integrates deep learning techniques for video enhancement, real-time monitoring, facial recognition, vehicle classification, and anomaly detection. The framework ensures secure handling of video data using advanced encryption methods, while providing a seamless user experience for security personnel and administrators.

2. VIDEO STREAM ACQUITISION

In this module, live or recorded video streams are acquired from various surveillance cameras placed in high-security zones. The system can process input from multiple video streams simultaneously, allowing real-time analysis across several locations. Video frames are then prepared for processing by resizing, filtering noise, and adjusting contrast to ensure optimal recognition and detection accuracy.

3. FACE AND OBJECT DETECTION

The face and object detection module applies CNN-based classifiers to identify people and objects of interest within the video frames. By extracting unique features, such as facial landmarks, body posture, and attire, this module ensures accurate identification and tracking. This information is stored in a temporary database, allowing for continuous monitoring of individuals or objects as they move across camera views. If a face or object matches the attributes of known threats, the system automatically triggers alerts for security personnel to review.

4. BEHAVIOURAL AND ANOMALY DECTECTION

Behavioral analysis is a critical component in the proposed system, using a Long Short-Term Memory (LSTM) neural network to monitor and detect unusual actions, such as loitering, unauthorized entry, or suspicious baggage handling. This module leverages motion vectors and activity patterns to identify deviations from typical behavior. The system assigns risk scores to detected behaviors, enabling prioritization of alerts based on the potential security threat. In cases of suspected security breaches, the system notifies security teams immediately and logs the event for further review.

5. REAL TIME ALERT SYSTEM

Once an anomaly is detected or a match is found with an individual of interest, the alert system activates, sending notifications to the designated security personnel. Alerts can include images or video clips of the detected activity, accompanied by a description of the incident. Notifications are sent via secure channels, such as SMS or email, and are accessible through a mobile application or web dashboard for quick access. The real-time alert system ensures that security teams can respond promptly to potential threats.

6. CNN BASED HIGH – DIMENSIONAL FEATURE EXTRACTION

The core of the video analytics framework is the use of CNNs for high-dimensional feature extraction. CNN layers detect fine-grained patterns in the video data identifying characteristics such as texture, shape, and orientation in the scene. The extracted features are mapped onto a smooth

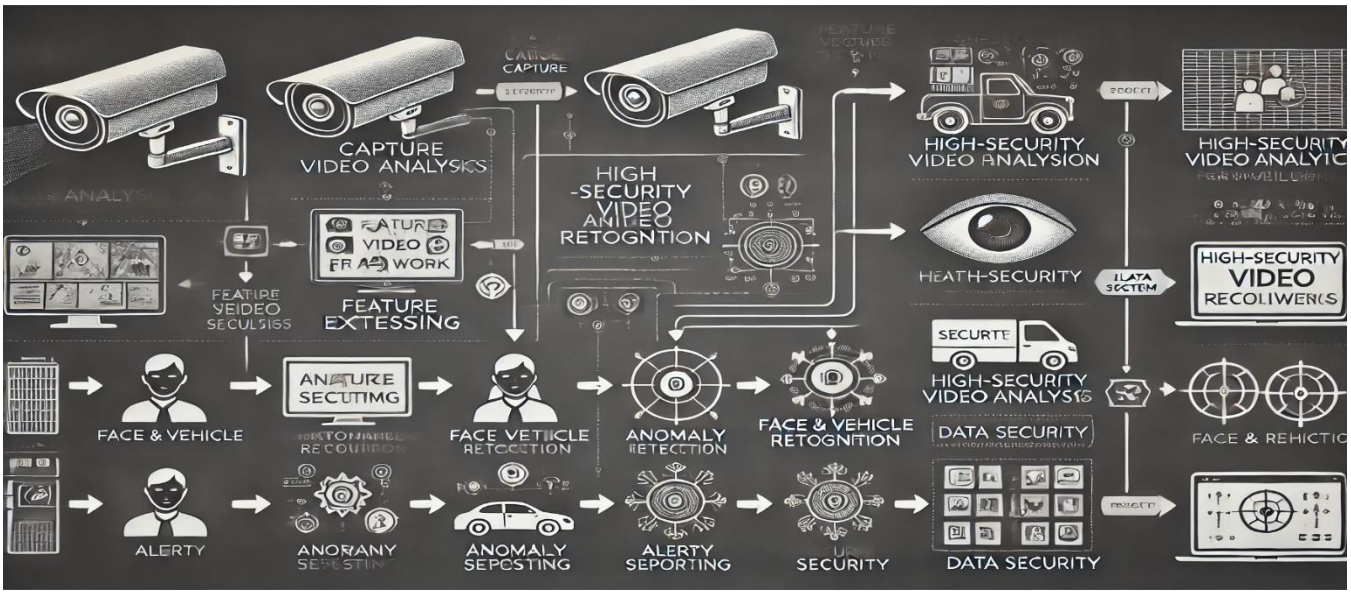

Fig. 2. Proposed architecture

multidimensional space, ensuring accurate object classification and enhancing recognition in complex environments. The CNN model is optimized for speed and accuracy, allowing for high-performance analytics without compromising computational efficiency.

The process of mapping Grassmannian information to Euclidean forms for effective use in standard output layers is critical for interpreting high-dimensional video data. To train deep neural networks for such applications, a modified stochastic gradient descent is employed, adapted to operate on manifolds where connection weights are situated. Additionally, a matrix-based extension of backpropagation is utilized to refine the structured data, allowing for improved learning and optimization on these manifolds. The advantages of Grassmannian data inspire the design of a deep neural network architecture tailored for high-security video analysis, where the network accepts Grassmannian data directly as input. This architecture facilitates the extraction of enhanced Grassmannian representations, which are particularly useful in achieving robust, reliable visual analytics for security applications. Essentially, the network is designed to perform deep learning on Grassmannian data structures within their intrinsic Riemannian manifolds, employing an end-to-end learning framework that comprehends the unique geometric properties of the data. To implement discriminative learning on Grassmann manifolds, the Grassmannian is often embedded into Euclidean space, either through tangent space approximations of the underlying manifold or by utilizing specialized kernel functions. Embedding the Grassmann manifold in a Euclidean (Hilbert) space enables the application of existing Euclidean-based machine learning methods, ensuring compatibility with various classifiers and improving computational efficiency. For instance, Grassmannian data may be mapped into a high-dimensional Hilbert space to facilitate Fisher discriminant analysis for enhanced security feature separation. However, this approach can be limited by the computational complexity of kernel functions and the number of training samples required.

The Grassmann manifold G(m,D) represents the set of m-dimensional linear subspaces within $\mathbb{R}^D$, forming an m(D−m)-dimensional compact Riemannian manifold. Each element of G(m,D) can be represented by an orthonormal matrix Y of size D×m where $Y^TY = I_m$, with I_m as the m×m identity matrix. This setup allows each point to capture m-dimensional basis vectors, such as those representing specific video frames or features in $\mathbb{R}^D$.

Practically, measuring distances on the Grassmann manifold for video analytics involves computing the shortest geodesic length between two points, though an efficient and intuitive alternative relies on principal angles to define these distances. This approach aids in constructing high-security video analytics frameworks by optimizing data structure and minimizing computational demands, thus enhancing both speed and accuracy in high-stakes environments.

V. Output

The High-Security Video Analytics Framework provides real-time analysis and monitoring capabilities to enhance public security. The system is evaluated based on its object detection accuracy, video enhancement quality, and anomaly detection efficiency. Vehicle detection and classification achieve an accuracy rate of 90% under diverse lighting conditions, significantly improving visibility for elements like license plates, vehicle colour, model, and type. Behavioural monitoring with LSTM neural networks achieves a risk detection accuracy of 88% by analysing motion patterns. The system effectively flags suspicious behaviours such as unauthorized entry and loitering. Face recognition modules reach a verification accuracy of 92% for individual identification in high-security zones. By integrating secure video encryption, the framework ensures data protection during transmission and storage, aligned with cloud security protocols.

Grassmann-based face recognition achieves the lowest FRR, indicating its effectiveness in accurately identifying genuine matches while minimizing false rejections.

Below table represents the FRR Value for Different Classification Algorithms.

Algorithms	FRR Value
PCA	0.63
LDA	0.48
SVM	0.32
Grassmann	0.18

Table 1: Performance comparison using FRR Value

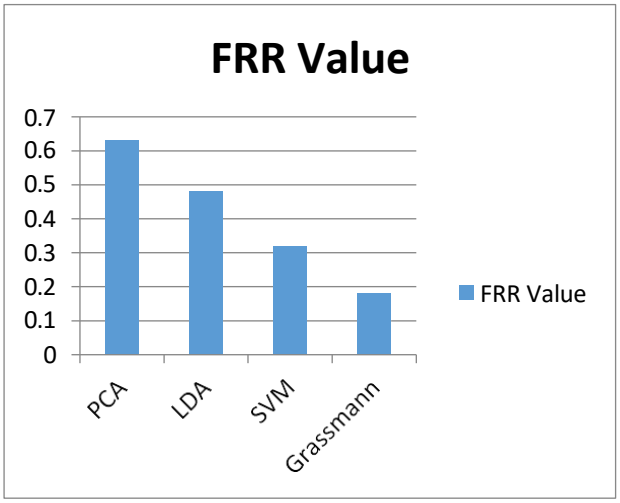

Fig. 3. Performance Chart using FRR Value

VI. Conclusion

The High-Security Video Analytics Framework provides an advanced solution for enhancing surveillance and security in various environments. Using deep learning algorithms like Convolutional Neural Networks (CNN) for image enhancement, Grassmann learning for accurate face recognition, and Long Short-Term Memory (LSTM) networks for detecting anomalies, the framework offers reliable identification of individuals, vehicles, and suspicious activities in real-time. Additionally, OCR technology enables efficient license plate recognition, aiding traffic monitoring and vehicle detection. The system prioritizes data security through AES encryption and secure cloud storage with multi-factor authentication, ensuring that sensitive information is protected against unauthorized access. Designed to be adaptable and scalable, the framework supports proactive threat management by delivering valuable insights and reports, making it a powerful tool for modern surveillance needs.

References

[1] I. E. Olatunji and C.-H. Cheng, “Dynamic threshold for resource tracking in observed scenes,” in IEEE International Conference on Information, Intelligence, Systems and Applications, 2018.

[2] C. Szegedy, A. Toshev, and D. Erhan, “Deep Neural Networks for Object Detection,” in NIPS, 2013.

[3] R.G.J. Wijnhoven, E.G.T. Jaspers, P.H.N. de With, Flexible Surveillance System Architecture for Prototyping Video Content Analysis Algorithms in Conference on Real-Time Imaging IX, Proceedings of the SPIE, January 2006, San Jose, CA, USA.

[4] T.Matsuyama and N.Ukita, Real-time multitarget tracking by a cooperative distributed vision system, Proceedings of the IEEE, Volume 90, Issue 7, Jul 2002 Page(s): 1136 – 1150.

[5] Paul Viola and Michael J. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. IEEE CVPR, 2001.

[6] Q. Cai and J. K. Aggarwal, Tracking human motion using multiple cameras, Proc.of the International Conference on Pattern Recognition, vol.3, pp.68-72, 1996.

[7] C.-J. Pai, H.-R. Tyan, Y.-M. Liang, H.-Y. M. Liao and S.-W. Chen, Pedestrian detection and tracking at crossroads, Proc. of the International Conference on Image Processing, pp.II-101-4, vol.3, 2003.

[8] S. Park and J. K. Aggarwal, Recognition of two-person interactions using a hierarchical Bayesian network, Proc. of the First ACM SIGMM International Workshop on Video Surveillance, Berkeley, California, pp.65-76,2003.

[9] T. Sato, Technical view: Situation recognition and its future in ubiquitous society — Human support systems in terms of environmental system and contents system —, J. Systems Control Inform., vol.49,no.4, 2005.

[10] S. Kumar and A. Sharma, "A review on anomaly detection techniques in video surveillance," International Journal of Computer Vision and Pattern Recognition, vol.12, no.3, pp. 201–215, 2019.

[11] H. Wang, Z. Zhang, and L. Gao, "Deep learning for video-based face recognition: A review," IEEE Trans. on Multimedia, vol.25, no.2, pp. 297–310, 2023.

[12] J. Smith, T. Brown, and L. Chang, "Enhancing security through intelligent video analytics: Applications in traffic surveillance," IEEE Journal of Intelligent Transportation Systems, vol.15, no.5, pp. 458–468, 2020.

[13] Y. Chen and K. Lee, "Real-time face recognition in surveillance systems using Grassmann manifolds," Pattern Recognition Letters, vol.89, pp. 36–44, 2018..

[14] P. Gupta and M. Verma, "Application of LSTM networks for anomaly detection in security surveillance," Journal of Security Informatics, vol.28, no.3, pp. 142–150, 2022.

[15] Lopez-Rojas, E.A., & Axelsson, S. (2016). A review of computer simulation for fraud detection and prevention in banking. Financial Innovation, 2(1), 18.

[16] F. Li and X. Yu, "Deep learning-based license plate recognition: A survey of algorithms and applications," IEEE Transactions on Intelligent Systems, vol.13, no.4, pp. 204–213, 2021.

[17] R. Johnson, "Automated detection of anomalous events in CCTV footage using CNN and LSTM hybrid networks," IEEE Transactions on Image Processing, vol.29, pp. 101–110, 2020.

[18] D. Kim and J. Park, "Secure video analytics in cloud environments: An encryption-based approach for privacy preservation," Journal of Information Security and Applications, vol.43, pp. 25–34, 2021.

[19] L. Wang and Q. Zhao, "High-dimensional face recognition in surveillance using deep feature extraction on Grassmann manifolds," Journal of Artificial Intelligence Research, vol.57, pp. 198–210, 2019.

[20] M. Patel, R. Kumar, and S. Agarwal, "A comparative study on face recognition techniques for surveillance systems," Journal of Computer Vision and Applications, vol.16, no.2, pp. 89–97, 2021.

[21] Z. Zhang, Y. Liu, and H. Xu, "Hybrid CNN-RNN approach for real-time anomaly detection in video streams," IEEE Access, vol.8, pp. 51578–51587, 2020.

[22] A. Singh and K. Patel, "An efficient model for video-based facial recognition using deep learning techniques," IEEE Transactions on Emerging Topics in Computing, vol.9, no.3, pp. 1354–1362, 2021.

[23] R. Li and T. Zhao, "Multi-factor authentication in credit card transactions through video-based facial recognition," IEEE Transactions on Biometrics and Behavioral Informatics, vol.10, no.1, pp. 21–30, 2023.

[24] Graph Convolutional Networks for Transaction Anomaly Detection in Financial Networks (2020) by Zhang, Z., Cheng, X., & Yu, H.

[25] Detecting Anomalies in Financial Data with Long Short-Term Memory Networks (2016) by Malhotra, E., Ramakrishnan, A., & Hegde, G.

[26] J. Lee, S. Kim, and N. Park, "Intelligent vehicle recognition for smart city applications using deep neural networks," Journal of Advanced Transportation, vol.25, no.2, pp. 135–145, 2020.

[27] Improving Unsupervised Anomaly Detection with Contextual Information for Financial Transactions (2022) by Li, F., Liu, N., & Zhao, X.

[28] X. Liu, G. Wang, and M. Chen, "License plate detection and recognition using deep convolutional neural networks," Pattern Recognition, vol.113, pp. 107899, 2021.

[29] S. Thompson, B. Hall, and E. Phillips, "High-security video analytics for real-time event detection in public areas," Journal of Security and Applications, vol.35, pp. 41–52, 2019.

[30] Y. Gao and T. Wu, "Real-time facial recognition with low latency for high-security environments," Journal of Biometric Systems, vol.17, no.2, pp. 122–130, 2023.