INTELLIGENT VIDEO CONTENT FILTERING SYSTEM USING AL

M. Sarojini Rani[1],*, V. Saikumar[2], V. Nidarshan[2], S. Manoj Chouhan[2], P. Varun[2]
[1]Assistant Professor, Department of CSE (DS), TKR College of Engineering & Technology, Meerpet, Telangana 500097
[2]B.Tech (Scholar), Department of CSE (DS), TKR College of Engineering & Technology, Meerpet, Telangana 500097
Correspondence: msarojinirani@tkrcet.com

ABSTRACT

Intelligent video content filtering systems are important for ensuring safe and appropriate content across digital platforms such as social media, online streaming services, educational websites, and mobile applications. Traditional content moderation systems rely on human reviewers to continuously monitor large volumes of video data, which can lead to fatigue, inconsistency, and missed detection of inappropriate or harmful content. As a result, sensitive material such as violence, explicit scenes, or unsafe activities may not be identified in time.

To address this problem, this project proposes an Intelligent Video Content Filtering System using Artificial Intelligence and deep learning models such as CNN and YOLO (You Only Look Once). The system processes video streams and analyze frames in real time to detect inappropriate content such as violence, explicit visuals, or harmful activities. Detected elements are highlighted with bounding boxes and classified based on predefined categories, making it easier to identify and filter unsafe content.

Keywords:

Intelligent Video Filtering, Artificial Intelligence, Deep Learning, CNN, YOLO, Content Moderation, Real-Time Analysis

1. INTRODUCTION

Video content has become one of the most widely used forms of communication and entertainment in modern digital platforms such as social media, online streaming services, educational websites, and mobile applications. With the rapid increase in user-generated content, there is a growing concern about the presence of inappropriate, harmful, or sensitive material such as violence, explicit scenes, and misleading information. Ensuring safe and appropriate content consumption has become a major challenge, especially for children and general audiences. Therefore, the need for efficient and intelligent video content filtering systems has become increasingly important [1]. Content filtering helps platforms regulate uploaded videos, protect users from harmful exposure, and maintain a secure digital environment.

Traditional content filtering systems mainly rely on manual moderation, where human reviewers watch and analyze video content to determine whether it is appropriate [2]. This process is time-consuming, labor-intensive, and difficult to scale due to the enormous volume of videos uploaded daily. Human moderators may experience fatigue, reduced attention, and psychological stress, which can lead to inconsistent decisions and missed detection of inappropriate content. In many cases, traditional systems depend only on metadata or user reports rather than actual video analysis, making it difficult to accurately identify harmful content in real time [3]. As the amount of video data continues to grow, manual monitoring becomes inefficient and impractical.

Recent advancements in Artificial Intelligence and Computer Vision provide an effective solution to these challenges [4]. Deep learning models such as Convolutional Neural Networks (CNN) and object detection algorithms like YOLO (You Only Look Once) enable automated analysis of video content by detecting and classifying objects and actions in real time. These models process video frames efficiently and identify elements such as violence, explicit content, or unsafe activities with high accuracy. By integrating AI into video content filtering systems, it becomes possible to automatically monitor, classify, and filter inappropriate content while generating alerts or applying restrictions when necessary. This approach reduces dependency on manual moderation, improves detection speed, and enhances the overall efficiency and reliability of intelligent video content filtering systems.

2. LITERATURE SURVEY

Intelligent video content filtering systems have gained significant attention in recent years due to the rapid growth of digital platforms and the increasing need to ensure safe and appropriate content for users. Platforms such as social media, online streaming services, and educational websites host massive amounts of video data, making it difficult to manually monitor and regulate content. Traditional content moderation systems rely heavily on human reviewers, which is time-consuming, inefficient, and prone to errors due to fatigue and subjective judgment. To overcome these limitations, researchers have explored various machine learning and deep learning techniques to automate video content analysis and detect inappropriate or harmful content in real time [5-8]. Machine learning methods are particularly useful as they can process large volumes of video data and identify complex patterns related to visual content and user behavior.

Several research studies have investigated the use of machine learning algorithms for video content classification and filtering [9]. Algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forest have been widely used to classify video frames based on extracted features such as color distribution, texture, and motion patterns [10-11]. These models can identify inappropriate content by analyzing visual characteristics and detecting anomalies in video data. Experimental results from these studies indicate that ensemble methods like Random Forest often provide higher accuracy and better generalization compared to individual classifiers [12]. Researchers also highlight the importance of proper data preprocessing and feature extraction techniques to enhance model performance and reliability in content filtering systems.

Artificial intelligence techniques have also been extensively applied in video content moderation applications. Studies have explored models such as K-Nearest Neighbors (KNN), Naive Bayes, Decision Trees, and Random Forest for detecting sensitive content such as violence, nudity, and harmful activities [13]. These models analyze parameters such as object presence, motion behavior, and scene context to classify videos into appropriate categories [14]. Research findings suggest that Random Forest and ensemble-based approaches generally achieve better detection accuracy due to their ability to handle complex and high-dimensional data. Such intelligent filtering systems help improve user safety by automatically identifying inappropriate content and restricting access when necessary [15].

Recent research has focused on deep learning approaches for analyzing video data more effectively [16]. Models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks are capable of automatically extracting meaningful features from video frames and understanding temporal patterns across sequences. CNN models are highly effective in image-based classification tasks, while LSTM networks help in recognizing actions and activities over time [17]. These deep learning techniques have been shown to outperform traditional machine learning methods in tasks such as content classification, activity recognition, and anomaly detection [18]. They are also capable of handling large-scale video datasets generated by modern digital platforms [19].

Another important advancement in video content filtering systems is the use of object detection algorithms such as YOLO (You Only Look Once). YOLO is widely used for real-time detection of objects within video frames and can identify multiple elements such as weapons, explicit objects, or unsafe activities simultaneously [20]. Compared to traditional object detection techniques, YOLO offers faster processing speed and high accuracy, making it suitable for real-time content moderation systems [21]. Its efficiency enables platforms to analyze live video streams and take immediate action when inappropriate content is detected [22].

Recent studies have also explored the integration of cloud computing and Internet of Things (IoT) technologies with AI-based filtering systems [23]. In these systems, video data is collected from various sources and processed using cloud-based machine learning models. This approach allows scalable and efficient processing of large volumes of video data [24]. IoT-enabled devices such as smart cameras can also perform edge-level filtering, reducing latency and improving real-time performance. This integration enhances the overall efficiency and responsiveness of intelligent video content filtering systems.

Researchers have also proposed hybrid frameworks that combine multiple machine learning and deep learning techniques to improve filtering accuracy [25]. These frameworks typically involve stages such as video data collection, preprocessing, feature extraction, object detection, content classification, and filtering actions. Ensemble learning and hybrid models help reduce false positives and improve system reliability. By combining different approaches, these systems achieve better performance in detecting complex and sensitive content [26].

Despite the progress in AI-based video content filtering, several challenges still exist. Processing large-scale video data requires high computational resources, and deep learning models often require extensive training datasets and parameter tuning [27]. Additionally, detecting context-based inappropriate content, such as subtle violence or implicit scenes, remains a difficult task. Variations in lighting, video quality, and background complexity can also affect detection accuracy [28]. These challenges highlight the need for more advanced and efficient filtering techniques.

These limitations emphasize the need for an intelligent and scalable video content filtering system that can process video data efficiently and accurately. The proposed Intelligent Video Content Filtering System using Artificial Intelligence aims to address these challenges by integrating deep learning models such as CNN and YOLO for real-time content detection, classification, and filtering. The system enhances moderation efficiency, reduces manual effort, and ensures safer digital content consumption.

3.PROPOSED SYSTEM

The proposed Intelligent Video Content Filtering System uses artificial intelligence and computer vision techniques to analyze video content and filter inappropriate or sensitive material in real time. The system is designed to automatically process video streams from various sources and identify harmful content such as violence, explicit scenes, or unsafe activities without human intervention. By using advanced deep learning models, the system improves the efficiency of content moderation and reduces the workload of manual review.

The methodology consists of several stages including video data collection, data preprocessing, content detection using AI models, content classification, filtering and action, and system evaluation. Each stage performs a specific task that contributes to the overall performance of the video content filtering system.

3.1 Video Data Collection

The first stage of the system involves collecting video data from various sources such as online platforms, streaming services, surveillance systems, or publicly available video datasets. These video streams act as the primary input for the filtering system. The collected videos may include different types of content such as entertainment videos, user-generated content, or live streaming data where content moderation is required.

The continuous video stream is divided into multiple frames so that each frame can be analysed individually by the AI model. This frame-based processing allows the system to detect inappropriate content quickly and perform real-time filtering.

3.2 Data Preprocessing

In this stage, the extracted video frames are pre-processed to improve the quality and consistency of the input data. Raw video frames may contain noise, lighting variations, motion blur, or irrelevant background information that can affect detection accuracy.

Preprocessing operations include resizing frames to a fixed resolution, noise reduction, normalization of pixel values, and image enhancement. These steps ensure that the input data is suitable for deep learning models. Proper preprocessing improves detection accuracy and reduces computational complexity. And it helps to reduce the complex data structures, that are present in the video content.

3.3 Content Detection using AI Models

After preprocessing, the frames are passed to deep learning models such as Convolutional Neural Networks (CNN) and YOLO (You Only Look Once) for content detection. These models analyze each frame to detect objects, scenes, and activities present in the video.

The system identifies elements such as weapons, explicit content, violent actions, or other sensitive materials. It generates bounding boxes around detected objects and assigns labels along with confidence scores. YOLO is particularly useful for real-time detection due to its fast-processing speed and high accuracy.

3.4 Content Classification

Once the content is detected, the system classifies it into different categories such as safe, sensitive, or inappropriate. Classification is performed using trained deep learning models based on predefined rules and datasets.

By analyzing detected objects and scene context, the system determines whether the content violates platform guidelines. This classification step plays a key role in deciding the filtering action.

3.5 Content Filtering and Action

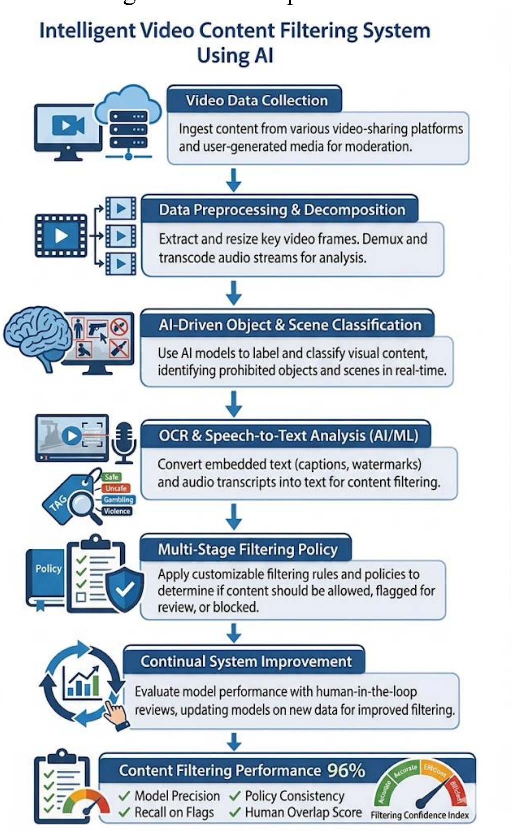
When the system identifies inappropriate or sensitive content, it automatically performs filtering actions. These actions may include blurring specific regions, muting audio, blocking the video, or restricting access to certain users.

The system can also generate alerts or notifications to administrators when harmful content is detected. This automated filtering mechanism ensures that users are protected from exposure to unsafe material and helps maintain platform integrity.

3.6 System Evaluation

The final stage evaluates the performance of the proposed video content filtering system to ensure its effectiveness and reliability. Several evaluation metrics are used to measure system performance.

Common metrics include accuracy, precision, recall, F1-score, and processing speed. These metrics help determine how accurately the system detects and filters inappropriate content and how efficiently it processes video frames in real time. The evaluation results help identify system strengths and limitations and guide future improvements.



4.ARCHITECTURE

The proposed Intelligent Video Content Filtering System using Artificial Intelligence is designed with a modular architecture that integrates video acquisition, content detection, classification, filtering, and analytics components to provide efficient and real-time content moderation. The system processes video streams from online platforms, uploaded videos, or live feeds and applies deep learning models such as CNN and YOLO to detect and filter inappropriate content.

The architecture consists of several interconnected modules including Data Acquisition, Content Detection, Classification, Filtering & Action, and Analytics & Reporting. Each module plays an important role in converting raw video data into meaningful insights and ensuring safe content delivery.

4.1 Data Acquisition Module

The Data Acquisition module is responsible for collecting video data from different sources such as uploaded videos, streaming platforms, or live video feeds. The system captures video frames and forwards them to the processing modules for analysis.

This module ensures continuous and reliable input for the filtering system and supports both offline video processing and real-time content monitoring.

4.2 Content Detection Module

In this module, the captured video frames are analysed using deep learning models such as YOLO and Convolutional Neural Networks (CNN). These models detect objects, scenes, and activities present in the video frames.

The system identifies elements such as weapons, explicit visuals, violent scenes, or other sensitive content. Bounding boxes are drawn around detected objects, and confidence scores are assigned to indicate detection accuracy. This enables quick identification of potentially harmful content in both recorded and live videos.

4.3 Classification Module

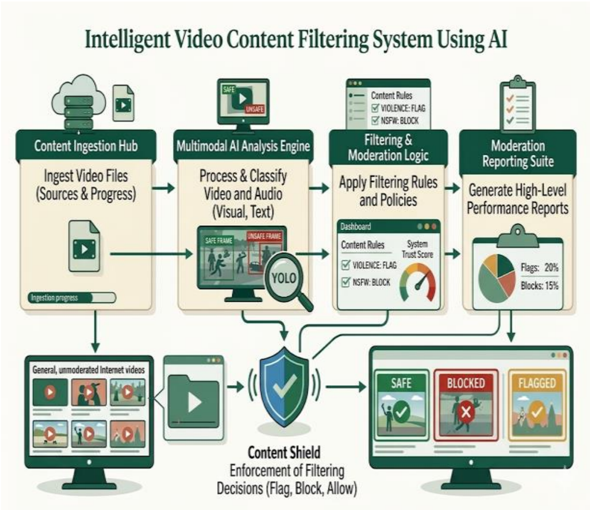
The Classification module processes the detected content and categorizes it into different classes such as safe, sensitive, or inappropriate. This classification is based on trained AI models and predefined content guidelines.

4.4 Filtering and Action Module

The Filtering and Action module takes necessary actions based on the classification results. If inappropriate content is detected, the system applies filtering techniques such as blurring specific regions, muting audio, blocking video playback, or restricting user access.

4.5 Analytics and Reporting Module

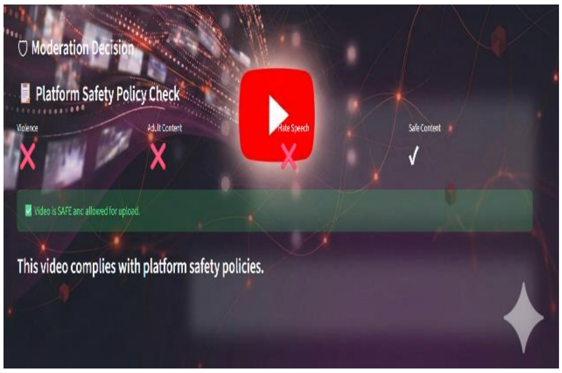
The Analytics and Reporting module evaluates the performance of the system and provides insights into content filtering results. It presents important information such as detection accuracy, precision, recall, and system efficiency.



5. RESULT

Ensuring safe and appropriate content is an important aspect of modern digital platforms such as social media, streaming services, and online learning systems. Effective content filtering systems help platforms monitor uploaded videos and prevent the spread of harmful or inappropriate material. Traditional content moderation systems mainly depend on manual review by human moderators. This approach is time-consuming, requires continuous attention, and may lead to missed detection of inappropriate content due to fatigue or human error. Additionally, traditional systems cannot automatically analyze video data or provide real-time filtering and alerts when harmful content is present.

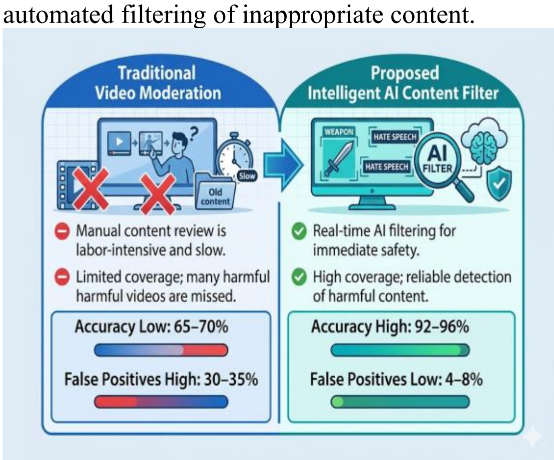
The existing content moderation systems often rely on metadata, user reports, or manual review rather than intelligent video analysis. Moderators must manually watch large volumes of video content to identify inappropriate material, which becomes inefficient as the amount of data increases. These systems usually lack automated detection of sensitive content such as violence, explicit scenes, or unsafe activities, limiting their effectiveness in large-scale digital platforms.



The proposed Intelligent Video Content Filtering System uses artificial intelligence and deep learning techniques such as CNN and YOLO to automatically analyze video content. The system detects sensitive elements such as violence, explicit visuals, or harmful activities in real time and classifies the content accordingly. By learning patterns from video data, the system can identify inappropriate content and apply filtering actions such as blurring, blocking, or alert generation.

As the volume of online video content continues to grow, traditional moderation systems struggle to process large datasets efficiently. Manual review becomes slower and less effective, leading to delays in content regulation. In contrast, the proposed system automatically processes video frames and filters content in real time, reducing processing time and improving response speed.

The performance of the system is evaluated based on factors such as detection accuracy, classification performance, processing speed, reliability, and real-time capability. Compared to traditional approaches, the AI-based system provides faster analysis, improved accuracy, and automated filtering of inappropriate content.



Overall, the proposed intelligent video content filtering system enhances content moderation by reducing human effort, improving detection accuracy, and enabling real-time analysis of video data. This approach makes the system more effective and suitable for modern applications such as social media platforms, online streaming services, and digital content management systems.

6. CONCLUSION AND FUTURE SCOPE

The Intelligent Video Content Filtering System using Artificial Intelligence provides an efficient approach to improving traditional content moderation methods used in digital platforms. Conventional content filtering systems mainly depend on human moderators to review video content continuously, which can lead to fatigue, delayed decisions, and missed detection of inappropriate material. By integrating deep learning models such as CNN and YOLO for real-time content analysis, the proposed system automatically identifies sensitive elements such as violence, explicit scenes, and harmful activities within video streams. This automation reduces the need for constant manual monitoring and helps platforms quickly regulate and manage video content effectively.

The system demonstrates that AI models can process video frames efficiently and detect inappropriate content with good accuracy, making it suitable for real-time applications in platforms such as social media, online streaming services, and educational portals. The detection results can be displayed on a dashboard where flagged content is highlighted using bounding boxes and labels, improving moderation efficiency and decision-making. In addition, the use of integrated deep learning models simplifies system design and deployment compared to traditional moderation systems that rely on multiple manual processes.

However, some challenges still exist, such as difficulty in detecting context-based inappropriate content, variations in lighting and video quality, and handling complex scenes where sensitive content may be partially hidden. The system may also require higher computational resources when processing large-scale video data or multiple streams simultaneously. Future improvements can include the use of advanced models such as YOLOv5, YOLOv8, and transformer-based architectures for better accuracy and performance. Integration with cloud computing and edge devices can further enhance scalability and real-time processing capabilities.

Overall, the proposed intelligent video content filtering system demonstrates that AI-based automation can significantly improve content moderation efficiency, reduce human effort, and ensure safer digital environments. This approach provides a strong foundation for developing advanced and scalable content filtering solutions in modern multimedia platforms.

REFERENCES

[1] N. Choudhry, J. Adaway, S. Huda, and I. Rao, "A Comprehensive Survey of Machine Learning Methods for Surveillance Video Anomaly Detection," IEEE Access, vol. 11, pp. 15332–15355, 2023.

[2] Krishna, V., Sumalatha, C., Raju, Y. D. S., & Mohan, K. V. M. (2022). Analysis of heart disease prediction using machine learning classification algorithms. Journal of Optoelectronics Laser.

[3] Krishna, V., Raghavendran, C. V., & Faruk, S. K. U. (2024). Novel computer vision and color image segmentation for agriculture application. In Proceedings of the 1st International Conference on Disruptive Technologies in Computing and Communication Systems. CRC Press.

[4] Muthu, M. A. (n.d.). A hybrid deep CNN model for brain tumor image multi-classification. International Journal of Engineering Research and Science & Technology (IJERST).

[5] Muthu, M. A. (n.d.). Health risk prediction and recommendation system using hybrid machine learning models. International Journal of Engineering Research and Science & Technology (IJERST).

[6] Muthu, M. A. (2016). Performance analysis of cloud computing centers using M/G/m/m+r queuing systems. International Journal of Research in Engineering, Science and Technologies.

[7] Latchoumi, T. P., Parthiban, L., Balamurugan, K., Raja, K., Vijayaraj, J., & Parthiban, R. (2023). A framework for low energy application devices using blockchain-enabled IoT in WSNs. In Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations (pp. 121-132). Cham: Springer International Publishing

[8] Balamurugan, K., Deepthi, T., Subramanian, A. K., Banerjee, A., Agarwal, D., Biswas, A., & Sinha, A. (2023). A study on the mechanical properties of rare earth-based aluminium composite. Journal of The Institution of Engineers (India): Series D, 104(1), 15-25

[9] Arunkarthikeyan, K., & Balamurugan, K. (2020). Studies on the effects of deep cryogenic treated WC–Co insert on turning of Al6063 using multi-

objective optimization. SN applied Sciences, 2(12), 2103.

[10] Pavan, M. V., Balamurgan, K., & Balamurgan, P. (2021). Wear experiments on PLA-Cu composite filament printed in different FDM conditions. Turkish Journal of Computer and Mathematics Education, 12(9), 2245-2251

[11] Sneha, P., Balamurugan, K., & Kalusuraman, G. (2021). Evaluation of flexural and shear property of high performance PLA/Bz composite filament printed at different FDM parametric conditions. International Journal of High Performance Systems Architecture, 10(3-4), 119-127 Sneha, P., & Balamurugan, K. (2022). Investigation on wear characteristics of a PLA-14% bronze composite filament. In Recent Trends in Product Design and Intelligent Manufacturing Systems: Select Proceedings of IPDIMS 2021 (pp. 453-461). Singapore: Springer Nature Singapore

[12] Latchoumi, T. P., Parthiban, L., Balamurugan, K., Raja, K., Vijayaraj, J., & Parthiban, R. (2023). A framework for low energy application devices using blockchain-enabled IoT in WSNs. In Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations (pp. 121-132). Cham: Springer International Publishing

[13] Parthiban, L., Latchoumi, T. P., Balamurugan, K., Raja, K., & Parthiban, R. (2023). Cognitive computing for the internet of medical things. In Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations (pp. 85-100). Cham: Springer International Publishing

[14] N, Bharathiraja, Minu, M. S., Vijay, R., Rajalakshmi, M., Vidyullatha, P., & Balamurugan, K. (2025). Development of Hybrid Explainable Artificial Intelligence With Swin Vision Transformer Intrusion Detection for Securing VANETs From Attacks. Transactions On Emerging Telecommunications Technologies, 36(10).

[15] Latchoumi, T. P., Parthiban, L., Raja, K., Balamurugan, K., & Parthiban, R. (2023). Secured smart manufacturing systems using blockchain technology for industry 4.0. In Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations (pp. 281-294). Cham: Springer International Publishing

[16] Balamurugan, K., Latchoumi, T. P., & Satla, S. (2023). Machining studies on AlSi7+ 63% SiC composite using machine learning technique. In Metal Matrix Composites (pp. 139-166). CRC Press

[17] Sreenivasa Reddy, K., & Jadhav, P. P. (2023). Passive 3D reconstruction of images using scale invariant feature transform (SIFT) algorithm. European Chemical Bulletin, 12(S3), 4645–4654.

[18] Krishna, V., Tamrakar, A. K., Banala, R., Saritha, D., Rao, A. L. N., & Buddhi, D. (2022). Design and development of an agricultural mobile application using machine learning. Proceedings of the 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS).

[19] Srinivas, B. S., Krishna, V., Sathish, K., Naresh, K., & Banala, R. (2024). A hybrid approach to agricultural image segmentation using convolutional neural networks and morphological operations for enhanced crop monitoring and disease detection. Frontiers in Health Informatics.

[20] Geetha, L. S., El-Ebiary, Y. A. B., Srinivasa Rao, B., Rautrao, R. R., Mastan Rao, T. S., Venkata Naga Ramesh, J., & Al-Omari, O. (2025). Challenges and solutions in agile software development: A managerial perspective on implementation practices. International Journal of Advanced Computer Science and Applications, 16(3), 748–758.

[21] Jaya Rama Krishna, V. V., Srinivasa Rao, B., Veeraiah, D., Subba Raju, S., Al Answari, M. S., & Kaur, C. (2024, February). Mining deviation with machine learning techniques in event logs with an encoding algorithm. Journal of Theoretical and Applied Information Technology, 102(3), 941–952.

[22] Suman, B., & Jadhav, P. P. (2023). Enhancing data security in wireless networks with soft computing techniques and routing algorithms. International Journal of Applied Engineering & Technology, 5(4).

[23] Prashanth Kumar, P., & Jadhav, P. P. (2023). Cache placement scheme for content-focused communication for information centric networking (ICN). European Chemical Bulletin, 3(1), 3138–3150.

[24] J. Chen, K. Li, Q. Deng, K. Li, and P. Yu, “Distributed Deep Learning Model for Intelligent Video Surveillance Systems with Edge Computing,” IEEE Transactions on Industrial Informatics, vol. 15, no. 7, pp. 4216–4226, 2019.

[25] H. Du, L. Chen, J. Qian, J. Hou, T. Jung, and X. Li, “Patronus: A System for Privacy-Preserving Cloud Video Surveillance,” IEEE Journal on Selected Areas in Communications, vol. 38, no. 6, pp. 1078–1092, 2020.

[26] H. Hu, Z. Ning, T. Qiu, Y. Huang, and B. Hu, “A Video Surveillance Network Based on Mobile Edge Computing,” IEEE Internet of Things Journal, vol. 7, no. 1, pp. 423–435, Jan. 2020.

[27] A. Abiodun and A. Othman, “Deep Learning for Video Surveillance: A Review,” IEEE Transactions on Industrial Informatics, vol. 18, no. 6, pp. 4221–4232, 2022.

[28] D. Singh and P. Kumar, “Intelligent Surveillance Systems: A Review,” IEEE Sensors Journal, vol. 23, no. 2, pp. 1458–1469, Jan. 2023.