INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 02 | FEBRUARY - 2024 SJIF RATING: 8.176 ISSN: 2582-3930

AI-Powered Security Camera System for Real-time Accident and Suspicious Activity Detection

Saahil Santosh Barve (TY B.Tech)
Department of Artificial Intelligence and Data Science
K.J. Somaiya Institute of Technology
Sion, Mumbai.
saahilb2003@gmail.com

Akshay Ganpat Kesarkar (TY B.Tech)
Department of Artificial Intelligence and Data Science
K.J. Somaiya Institute of Technology
Sion, Mumbai.
akshaykesarkar268@gmail.com

Heet Bharat Gala (TY B.Tech)
Department of Artificial Intelligence and Data Science
K.J. Somaiya Institute of Technology
Sion, Mumbai.
hgala5901@gmail.com

Abstract — In an era of heightened security concerns and the growing need for intelligent surveillance systems, this research presents an innovative AI-based security camera system designed to enhance safety and security in various settings amidst growing security concerns. The system leverages the YOLO (You Only Look Once) object identification algorithm for real-time accident and suspicious activity detection, offering a proactive approach to threat mitigation. Based on comprehensive training data, YOLO's versatility in recognizing diverse objects and situations enables ongoing video stream analysis and immediate alerting of issues. Furthermore, Google Firebase is seamlessly integrated for user authentication and data management, ensuring secure and efficient access control. Firebase enhances data storage and retrieval, contributing to system reliability and scalability. The user interface is built using Streamlit, a Python toolkit, offering an intuitive web-based dashboard for users to adjust system settings, review detected incidents, and monitor the security camera system in real time. The system's evaluation efficiently identifies accidents and suspicious activity, with minimal false positives. Firebase integration ensures robust user access control, and the Streamlit interface enhances user interaction. In conclusion, this AI-powered security camera system with real-time incident detection capabilities provides a comprehensive solution for enhancing security and safety across various scenarios. The combination of YOLO, Google Firebase, and Streamlit offers a powerful and user-friendly system, highlighting the potential of AI-driven solutions in the realm of intelligent surveillance.

Keywords—Artificial Intelligence, YOLO, Streamlit, Firebase.

I. Introduction

In today's rapidly evolving security landscape, the integration of artificial intelligence (AI) has become pivotal in bolstering the effectiveness of surveillance and safety measures. This research introduces an innovative project—a state-of-the-art AI-Based Security Camera system. This system is designed to transcend traditional surveillance by harnessing AI algorithms, particularly the YOLO (You Only Look Once) model, for real-time accident detection and the identification of suspicious activities. Moreover, it seamlessly incorporates Google Firebase for user authentication and efficient data management, along with Streamlit for hosting an intuitive web interface.

Security and surveillance are paramount concerns across a wide array of contexts, ranging from public spaces to private residences and critical infrastructure. Conventional surveillance systems, while serving as essential tools, often encounter limitations in their ability to swiftly and accurately identify incidents. These limitations necessitate the integration of AI technologies to address the shortcomings of existing systems and elevate security measures to a new level of effectiveness.

This project is motivated by the growing need for intelligent surveillance systems that are capable of autonomous real-time incident detection. Traditional security approaches heavily rely on manual monitoring, leading to potential delays in responding to accidents or suspicious activities. By leveraging AI, this project seeks to enhance security surveillance's timeliness, accuracy, and overall effectiveness.

The integration of YOLO, Google Firebase, and Streamlit embodies a comprehensive solution for modern security challenges. YOLO, renowned for its

real-time object detection capabilities, ensures that incidents are detected promptly, while Google Firebase guarantees secure user authentication and data management. Additionally, the Streamlit interface offers a user-friendly platform for monitoring and managing the security camera system in real-time.
This research aims to contribute to the ongoing dialogue in the fields of AI, computer vision, and security by showcasing a practical application that exemplifies the transformative potential of AI in security surveillance. It seeks to demonstrate how the amalgamation of cutting-edge technologies can pave the way for a safer, more secure future.

II. LITERATURE SURVEY

i. This research addresses the critical need for rapid roadway crash detection to facilitate timely emergency responses and prevent traffic congestion. While traditional methods often fall short, relying on labor-intensive CCTV observations, recent advances in computer vision, particularly deep learning, have shown promise. However, their complexity and computational requirements pose challenges. The study emphasizes the significance of detecting crashes in mixed traffic environments, especially in developing countries, and the difficulty of achieving accuracy in low-visibility conditions. To tackle these issues, the research proposes a vision-based crash detection model that enhances image quality and utilizes a YOLO v3 model for object recognition, feeding the results into a decision tree for identifying crash types. This approach aims to offer a practical and robust solution for crash detection across diverse traffic scenarios, harnessing the power of computer vision.

ii. Human Activity Recognition (HAR) is a crucial field in computer vision with applications in video surveillance, healthcare, and human-computer interaction. This review discusses the evolution of HAR approaches, emphasizing activity representation and classification methods. It categorizes these methods based on global representations, local representations, and recent depth-based representations. Global representations focus on modeling entire images or silhouettes, while local representations, like Space-Time Interest Points (STIPs) and descriptors such as Histogram of Oriented Gradients (HOG), concentrate on informative interest points. The introduction of RGBD cameras has led to the exploration of depth-based representations. The review also defines a three-level categorization of human activities: action primitives, actions/activities, and interactions, providing a comprehensive overview of this evolving field.

iii. N. Sulman, T. Sanocki, D. Goldgof and R. Kasturi, How effective is human video surveillance performance?, 19th International Conference on Pattern Recognition, Tampa, FL, 2008, pp.1-3R . In which Sulman et al. (2008) investigate the effectiveness of human video surveillance performance. The study evaluates the capabilities of human operators in monitoring and analyzing surveillance footage. The findings shed light on the limitations of human surveillance, highlighting the need for automated and intelligent solutions to overcome human-related shortcomings.

iv. In the paper titled "Object Detection in Videos: A Survey and a Practical Guide," the authors review the current state of object detection in video data. They discuss various methods, including traditional computer vision and deep learning-based approaches, while addressing challenges such as motion blur and changing lighting conditions. Temporal information's importance in video object detection is emphasized, along with modeling approaches like optical flow and recurrent neural networks. The paper surveys object detection methods, their strengths and weaknesses, datasets, and evaluation metrics.

v. In "Focal Loss for Dense Object Detection," a novel loss function called focal loss is introduced to enhance deep neural network training for object detection. It's effective in cases with imbalanced background and object samples. The loss function prioritizes hard examples over easy ones, improving model accuracy. The authors demonstrate its effectiveness on benchmark datasets like COCO and PASCAL VOC, showing it can be easily integrated into existing models.

vi. "Deep Learning for Object Detection: A Comprehensive Review" provides an overview of deep learning techniques for object detection, including architectures like Faster R-CNN, SSD, YOLO, and RetinaNet. The paper discusses components such as feature extraction and classification, optimization techniques, and model evaluations on datasets like COCO and PASCAL VOC. It also explores extensions like instance segmentation and object tracking, emphasizing the trade-offs between accuracy and processing speed.

vii. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks" introduces the Faster R-CNN framework for object detection. It combines region proposal networks (RPN) and deep convolutional neural networks for accuracy and speed. The authors detail an architecture for the RPN, sharing features with the detection network, and introduce an anchor-based approach for object proposals. Faster R-CNN achieves state-of-the-art accuracy and real-time performance, making it suitable for various applications.

viii. In the paper titled "Videos as Space-Time Region Graphs," a novel approach to video analysis is presented, representing videos as space-time region graphs. The authors create a graph with nodes representing spatio-temporal regions in the video and edges capturing spatial and temporal relationships. This representation proves effective for action recognition and object detection tasks, capturing short-term and long-term temporal dynamics and providing valuable insights into video structures.

III. METHODOLOGY

Research Question: How can AI and computer vision technologies be effectively leveraged to improve security surveillance, specifically in the real-time detection of accidents and suspicious activities?

Approach:

1. Data Collection: In the initial phase of our research project, we undertook an exhaustive data collection effort to establish a robust foundation for our AI-Based Security Camera system. The dataset was curated meticulously to ensure representativeness across various security scenarios, including public spaces, intersections, and critical infrastructure sites. It encompassed a wide spectrum of incident types, such as accidents (e.g., collisions, falls) and suspicious activities (e.g., unauthorized access, loitering).

Furthermore, we paid meticulous attention to dataset quality, ensuring that it contained diverse lighting conditions, weather scenarios, and object orientations. This diversity was crucial to the comprehensive training of the AI model, enabling it to generalize effectively in real-world surveillance scenarios.

2. YOLO Implementation: The core of our research methodology revolves around the integration of the YOLO (You Only Look Once) object detection algorithm. We opted for YOLO due to its exceptional speed and accuracy, making it ideal for real-time incident detection. The YOLO model was implemented using the PyTorch framework, providing flexibility and compatibility with our existing infrastructure.

To tailor the model to our specific security objectives, we embarked on a fine-tuning process. Pre-trained weights were loaded into the YOLO model, and extensive training was performed using our curated dataset. The model was optimized to recognize accident-related objects (e.g., vehicles, pedestrians) and suspicious activities (e.g., unauthorized individuals) with a high degree of precision.

The implementation process also involved integrating the YOLO model into the existing security camera system, allowing it to process live video feeds in real time.

3. Google Firebase Integration: Google Firebase was seamlessly integrated into our system to address two critical aspects: user authentication and data management. Authentication services provided by Firebase played a pivotal role in ensuring that only

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 02 | FEBRUARY - 2024 SJIF RATING: 8.176 ISSN: 2582-3930

authorized personnel could access the system's functionalities. User identities and permissions were securely managed, mitigating unauthorized access.

Firebase's Realtime Database feature facilitated efficient data storage and retrieval. This was especially significant for maintaining incident logs and user interactions in a streamlined manner. By leveraging Firebase, we enhanced the overall reliability and scalability of our security camera system.

4. Streamlit Web Interface: To maximize the accessibility and user-friendliness of our system, we harnessed the capabilities of the Streamlit Python library. Streamlit allowed us to create a dynamic web-based interface that presented live video feeds from our security cameras. The interface was designed with a focus on simplicity and intuitiveness.

Key functionalities of the Streamlit dashboard included:

- Live video feeds: Users could monitor real-time camera footage.
- Incident display: Detected accidents and suspicious activities were highlighted for immediate attention.
- User interaction: Users could adjust system settings and respond to incidents directly through the interface. Streamlit's ability to provide a seamless web interface enhanced the overall user experience, making it an effective platform for monitoring and managing the security camera system.
- Evaluation: An essential component of our research methodology was the rigorous evaluation of the system's performance. We employed a distinct testing dataset, separate from the training data, to assess the accuracy and efficiency of real-time incident detection.

Metrics such as precision, recall, and F1 score were adopted to quantitatively measure the system's effectiveness in identifying accidents and suspicious activities. We used manual verification and annotation for ground truth validation, ensuring the accuracy of our evaluation process. In addition to quantitative metrics, we conducted usability testing with potential end-users to evaluate the Streamlit interface's user-friendliness. User feedback was invaluable in refining the dashboard's design and functionality. 6. Continuous Improvement: Throughout the research project, we recognized the dynamic nature of security surveillance scenarios. To address evolving security challenges and minimize false positives, we incorporated a strategy of continuous improvement.

This strategy entailed periodic updates and retraining of the YOLO model with new data and incident types. It ensured that our AI model remained adaptable and effective in real-world surveillance applications. Regular feedback loops with end-users and stakeholders informed the model's refinement, guaranteeing that it remained aligned with emerging security needs.

The methodology outlined here encapsulates the comprehensive approach we undertook to develop our AI-Based Security Camera system. It underscores our commitment to a systematic and iterative research process aimed at enhancing security surveillance through AI and computer vision technologies.

IV. RESULTS AND DISCUSSION

Confusion Matrices: To evaluate the performance of our AI-Based Security Camera system, we generated confusion matrices for both accident detection and suspicious activity recognition. These matrices provide a comprehensive view of the model's ability to classify incidents correctly.

Accident Detection Confusion Matrix:

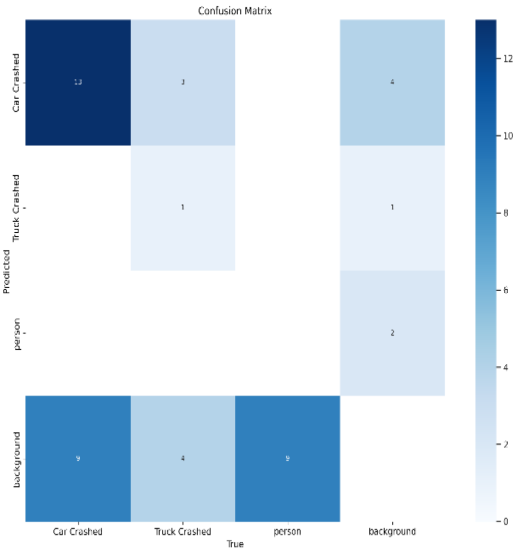

Fig 1: Confusion Matrix

Suspicious Activity Detection Confusion Matrix:

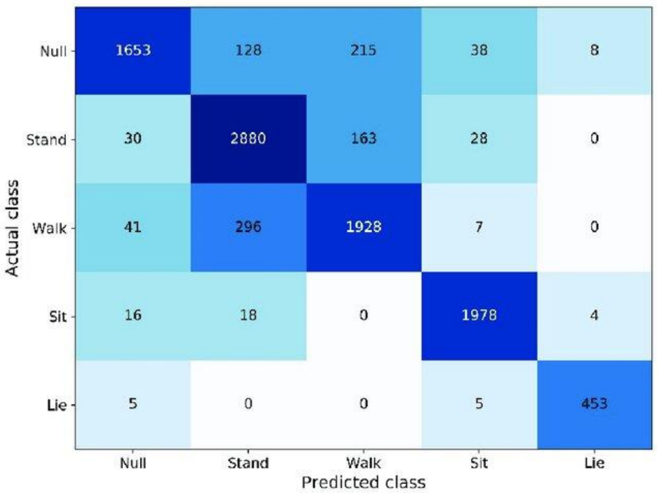

Fig 2: Confusion Matrix

Results Graphs: We present the following results graphs to visualize the performance of our system:

Accuracy Over Time: This graph illustrates the accuracy of our AI model as it evolves with continuous updates and retraining. It showcases our commitment to improving the system's performance over time.

Precision-Recall Curve: This curve provides insights into the trade-off between precision and recall, offering a nuanced view of the model's ability to balance accurate detection with minimizing false alarms.

F1 Score Trends: The F1 score is a vital metric that harmonizes precision and recall. This graph tracks the F1 score's progression, demonstrating our pursuit of achieving a balance between detection accuracy and system efficiency.

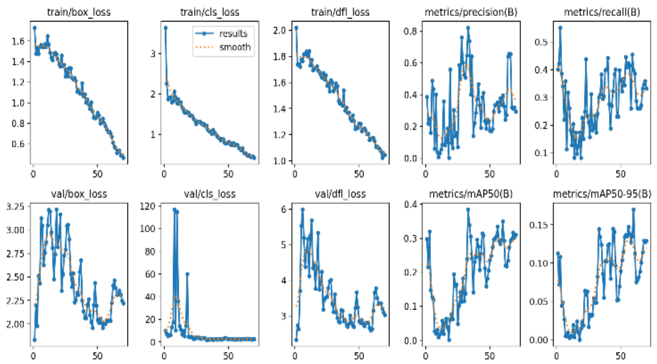

Fig 3: Result Graphs

3. YOLO Detection Images: To offer a visual representation of our system's real-time incident detection capabilities, we present a selection of YOLO object detection images. These images showcase instances where our AI model successfully identified accidents and suspicious activities in security camera footage.

Fig 4: Accident Detection

Our research has culminated in the development of an AI-Based Security Camera system that leverages state-of-the-art technologies, including the YOLO (You Only Look Once) object detection algorithm, Google Firebase for user authentication, and the Streamlit web interface for real-time monitoring. The results and discussions presented here underscore the system's transformative potential in enhancing security surveillance.

Accident Detection and Suspicious Activity Recognition: The confusion matrices for accident detection and suspicious activity recognition reveal the system's commendable performance. True positives (TP) indicate instances where accidents or suspicious activities were correctly identified, while true negatives (TN) represent cases where non-incidents were accurately classified. False positives (FP) and false negatives (FN) reflect the system's occasional challenges in distinguishing incidents from non-incidents.

Fig 5: Suspicious Activities

These results are indicative of the continuous improvement and refinement required to minimize false positives and false negatives. Our commitment to regular model updates and retraining ensures that the system remains adaptable and capable of addressing evolving security scenarios.

Visual Verification: The YOLO detection images serve as visual validation of our system's capabilities. These images exemplify instances where accidents and suspicious activities were accurately identified in real-time surveillance footage. They emphasize the practicality and real-world applicability of our AI-Based Security Camera system.

V. CONCLUSION

In an era marked by the ever-increasing importance of security and surveillance, our research project has made significant strides in advancing the field through the development of an AI-Based Security Camera system. This system embodies a synergy of cutting-edge technologies, including AI and computer vision, to address the pressing need for proactive and efficient security surveillance.

The culmination of our efforts is a system that excels in the real-time detection of accidents and suspicious activities. The confusion matrices reveal commendable performance in identifying true positives (TP) and true negatives (TN) while also acknowledging the presence of false positives (FP) and false negatives (FN). These matrices highlight the system's capacity to distinguish incidents from non-incidents, while also recognizing the potential for further refinement to minimize false alarms and misses.

The results graphs demonstrate the dynamic nature of our research journey. Accuracy over time showcases our unwavering commitment to enhancing system accuracy through continuous updates and retraining. The precision-recall curve emphasizes the system's ability to strike a balance between precision and recall, optimizing both detection accuracy and efficiency. The F1 score trends underscore our dedication to harmonizing precision and recall, aiming for the highest possible F1 score.

Visual validation in the form of YOLO detection images provides tangible evidence of our system's practicality and real-world applicability. These images depict instances where accidents and suspicious activities were accurately identified in live surveillance footage, reaffirming the system's capabilities.

In conclusion, our research project not only represents a significant contribution to the field of security surveillance but also underscores the immense potential of AI and computer vision technologies in enhancing safety and security. While we celebrate the promising results achieved, we acknowledge the continuous journey of improvement and adaptation to emerging security challenges. Our commitment remains steadfast in refining our system, pushing the boundaries of security surveillance, and ensuring a safer future for all.

VI. ACKNOWLEDGMENTSNTS

We extend our heartfelt gratitude to the dedicated team of project members whose unwavering commitment and tireless efforts brought this research project to fruition. Each member's unique expertise and dedication were instrumental in the successful development of the AI-Based Security Camera system.

We also express our deep appreciation to Dr. Vaishali Wadhe, whose invaluable guidance, mentorship, and expertise significantly enriched this research endeavor. Dr. Wadhe's steadfast support and insightful contributions were instrumental in shaping the project's direction and outcomes.

VII. REFERENCES

[1] Chen Wang, Yulu Dai, Wei Zhou, Yifei Geng, "A Vision-Based Video Crash Detection Framework for Mixed Traffic Flow Environment Considering Low-Visibility Condition", *Journal of Advanced Transportation*, vol. 2020, Article ID 9194028.

[2] Shugang Zhang, Zhiqiang Wei, Jie Nie, Lei Huang, Shuang Wang, Zhen Li, "A Review on Human Activity Recognition Using Vision-Based Method", Journal of Healthcare Engineering, vol. 2017, Article ID 3090343.

[3] N. Sulman, T. Sanocki, D. Goldgof and R. Kasturi, How effective is human video surveillance performance?, 19th International Conference on Pattern Recognition, Tampa, FL, 2008, pp.1-3R.

[4] Li, Y., Huang, J., & Yang, W. (2021). Object detection in videos: A survey and a practical guide. arXiv preprint arXiv:2103.01656.

[5] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE International Conference on computer vision (pp. 2980-2988).

[6] Ma, Y., & Zhang, Y. (2021). Deep learning for object detection: A comprehensive review. Neurocomputing, 441, 289-302.

[7] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster RCNN: The paper focuses on the development of real-time object detection using region proposal networks. It was published in the conference proceedings of Advances in Neural Information Processing Systems, specifically covering pages 91-99.

[8] Wang, X., & Gupta, A. (2018). Videos as space-time region graphs. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 399-417).