Impact Factor (SJIF): 6.126 E-ISSN: 2583-1615, P-ISSN: 3049-3498

International Journal of Research in Academic World

Received: 10/August/2025 IJRAW: 2025; 4(9):154-156 Accepted: 21/September/2025

Shoplifting Detection in Video Surveillance Using SlowFast

[1]Gagana TV and [*2]Asha S

[1]Department of Computer Science and Engineering, SCMS School of Engineering and Technology, Karukutty, Ernakulam, Kerala, India.

[*2]Assistant Professor, Department of Computer Science and Engineering, SCMS School of Engineering and Technology, Karukutty, Ernakulam, Kerala, India.

Abstract

Shoplifting is a major problem for retail stores, resulting in financial losses and operational inefficiencies. Traditional surveillance systems, while useful for monitoring activities, often require constant human supervision, making them prone to oversight and delays in response. To address this issue, a Shoplifting Detection System has been developed to automatically analyze surveillance videos and identify suspicious behavior. The system processes the uploaded video by extracting frames at regular intervals and analyzing them to detect the instances of shoplifting. When a shoplifting activity is detected, the system triggers an alarm, alerting security team to take immediate action. This approach enhances the security by reducing human error and ensuring quicker responses to such threats. The system follows a structured workflow where the video is first uploaded for analysis, followed by frame extraction and processing to identify unusual activities. When shoplifting is detected, the system suddenly activates the alarm to notify the security team. By automating the detection process, this system significantly reduces the need of manual monitoring, allowing security team to focus on other critical tasks. This not only improves overall store security but also helps prevent potential losses, making the monitoring process more efficient and reliable.

Keywords: Accuracy, SlowFast, DCSASS Dataset.

1. Introduction

Retail stores often face the challenge of preventing shoplifting, which results in substantial financial losses and operational inefficiencies. To mitigate these challenges, surveillance systems are commonly deployed to monitor activities in stores. However, manually observing multiple video feeds for prolonged periods is both time-consuming and prone to errors. To counter these limitations, a Shoplifting Detection System has been created to automatically scan surveillance videos and detect suspicious activity. The system processes the uploaded videos by extracting frames at regular intervals and analyzing them to detect potential instances of shoplifting. When such activity is identified, an audio alarm is triggered to alert security personnel for immediate intervention. This system enhances the effectiveness of surveillance and ensures quick responses, reducing the likelihood of undetected incidents. Variety

2. Related Work

The latest developments in human unusual activity detection using deep learning models and machine learning methods for identifying abnormal behaviors in video surveillance. Various studies show the capability of Deep Belief Networks (DBNs), Convolutional Neural Networks (CNNs), and ensemble models to provide excellent accuracy. Several models highlight the use of spatiotemporal feature extraction techniques and hybrid models for improving detection efficiency in real-world applications. The paper "Enhanced Human Activity Recognition Using Convolutional Neural Networks" [1] by Kajendran and Albert (2024) introduces a 2D Convolutional Neural Network (2D-CNN) combined with Support Vector Machine (SVM) to improve the accuracy of human activity recognition in video surveillance. The study highlights the importance of CNNs in extracting spatial features while SVM enhances classification performance. The model achieves an impressive accuracy of 95%, making it a robust approach for identifying suspicious behaviors. The paper "Suspicious Activity Detection Using Machine Learning" [2] by Kumbhar *et al.* (2023) presents the combination of Convolutional Neural Networks (CNN), Deep Belief Networks (DBN), and K-means clustering for detecting suspicious activities from video streams. The models combined provide improved anomaly detection with an accuracy of 90%. The study emphasizes the importance of preprocessing techniques such as background subtraction and motion estimation to improve model performance. "Automatic Unusual Activities Recognition" [3] by Ramzan *et al.* (2020) discusses the use of 2D and 3D CNN models for the detection of unusual activity in CCTV videos. While 2D CNN was at 77%, 3D CNN was 76%. The paper points out that while 3D CNN models are good at extracting spatiotemporal information, their computational expense may

*Corresponding Author: Asha S

restrict realtime use. The research proposes that 2D CNN models offer a computationally less demanding alternative with similar performance. Application for convenient diagnosis by patients and dermatologists, highlighting low-cost and non-invasive diagnostic techniques.

3. SlowFast Network

The SlowFast model is a two-stream convolutional neural network (CNN) designed to effectively capture both spatial and temporal information from video data. It overcomes the limitations of conventional models by utilizing two parallel pathways—Slow Pathway and Fast Pathway—to analyze video frames at different temporal resolutions.

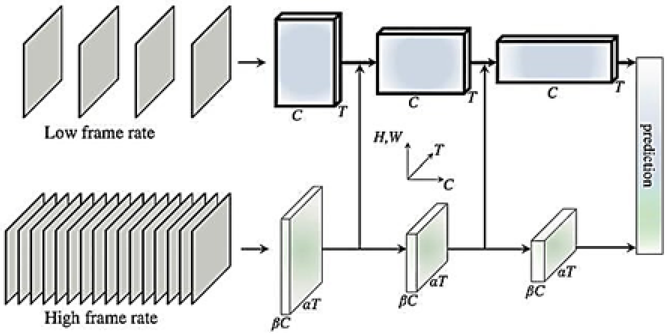

Fig 1: The SlowFast model

Input video frames are pulled out and resized to a standard resolution of 224×224 pixels in order to match the model. Frame normalization and augmentation techniques, including rotation and flipping, are applied to improve model generalization. Slow Pathway: Processes video frames at a low frame ratse, capturing detailed spatial details at information over longer intervals. It applies several convolutional blocks to get hierarchical features at higher depths:

- **Conv Block 1:** 64 channels
- **Conv Block 2:** 128 channels
- **Conv Block 3:** 256 channels

Fast Pathway: Processes video frames at a higher frame rate, capturing motion dynamics and short-term temporal variations. It uses fewer channels to maintain efficiency:

- **Conv Block 1:** 8 channels
- **Conv Block 2:** 16 channels
- **Conv Block 3:** 32 channels

The features extracted through both pathways are combined at the feature fusion stage, where information complementary to the slow and fast pathways is pooled to give a richer understanding of the video.
A fully connected classifier processes the fused features to predict whether the input video contains shoplifting activity.

4. Proposed System

The Shoplifting Detection System has a clearly defined process, beginning with video input from observation cameras or uploaded video. The videos are preprocessed, where frames are sampled at periodic intervals and resized to ensure uniformity. These frames are then analysed to identify suspicious activities that may indicate shoplifting. The system uses established rules and criteria to label the frames and decide whether there is any suspicious activity. On identifying possible shoplifting, the system sounds an audio alarm to notify security personnel, thus facilitating immediate action.
The framework backend is built utilizing Jar, which handles video transfer, forms the extricated outlines, and oversees the cautioning mechanism. Security officers can get to the framework through a web-based interface, where they can transfer recordings for preparing and be alarmed upon discovery of suspicious movement. The system is capable of handling multiple video feeds, making it suitable for deployment in large retail environments. By automating the detection process and providing real-time alerts, the system enhances security, reduces human workload, and minimizes the risk of undetected thefts.

5. Results and Discussion

The overall goal of this project was to develop an anti-shoplifting system for catching shoplifters in SlowFast model surveillance footage. The model takes input video frames and processes them through a two-stream network, with the slow path capturing fine spatial details and the fast path capturing motion dynamics to observe temporal variations.
The system was evaluated on a vast array of surveillance videos with typical behavior and instances of shoplifting. The SlowFast model achieved 89.42% accuracy in distinguishing shoplifting from regular behavior. The combination of spatial and temporal features allowed the system to identify suspicious activities effectively, leading to reliable performance.
Upon detecting shoplifting, the system generated an alarm and extracted the frame where the unusual activity occurred. The extracted frames were displayed for further analysis. The technique ensured that the system reacted as quickly and efficiently as possible when it sensed actual threats, as it is tailored for practical usage in surveillance contexts. The results validate that the proposed model can effectively detect suspicious behavior and trigger appropriate alerts in real time.

Table 1: Performance Evaluation Table

Metric	Class 0 (Normal)	Class 1 (Shoplifting)
Precision	0.97	0.97
Recall	0.99	0.86
F1-Score	0.98	0.91

Table 2: Comparison of Different Existing Algorithms and our Proposed Algorithm (AADSFR50)

Sr. No	Algorithm	Accuracy (%)
1	C3D [16]	23
2	TCNN [17]	28.4
3	3D Resnet 34 [18]	27.2
4	3D ConvNets [19]	45
5	Semi-supervised GAN [20]	40.9
6	VGG-16 [21]	72.66
7	VGG-19 [21]	71.66
8	FlowNet [22]	71.33
9	Our proposed Algorithm	89.42%

(a) Sample of Shoplifting Video Detected as Shoplifting. (b) Sample of Normal Video Detected as Normal.

Fig 2: Implementation Results of our Proposed Method

Conclusion

The proposed system successfully detects shoplifting activities in surveillance videos using the SlowFast model. The model processes input video frames through its two-stream structure, with the slow pathway capturing spatial information and the fast pathway dealing with motion dynamics. Through the integration of spatial and temporal features, the system efficiently detects suspicious activity and alerts an alarm upon detection of shoplifting. The extracted frame containing the unusual activity is displayed, enabling real-time decision-making and improving security monitoring. Compared to the research paper "Suspicious Activity Detection Using Machine Learning" by Bhagyashri Kumbhar *et al*. (2023), which used a Convolutional Neural Network (CNN) to identify abnormal behaviors, the SlowFast model shows better performance by capturing both spatial and temporal information. Whereas the CNN-based method focused on the analysis of static frames, the SlowFast model handles dynamic scenes efficiently and thus better suited for real-world surveillance.

References

1. Kajendran K & Albert J. Enhanced human activity recognition using convolutional neural networks. Expert Systems with Applications, 2024.
2. Kumbhar B, Pisal P, Kene K & Raut A. Suspicious activity detection using machine learning. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, May 2023.
3. Ramzan M, Abid A, Awan SM, Shahid M & Ahmad M. Automatic unusual activities recognition. Computers, Materials & Continua, 2020.
4. Bora TS & Rokade MD. Human suspicious activity detection system using CNN model for video surveillance. *International Journal of Advance Research and Innovative Ideas in Education (IJARIIE),* 2021.
5. Rodrigues R, Bhargava N, Velmurugan R & Chaudhuri S. Multi-timescale trajectory prediction for abnormal human activity detection. Winter Conference on Applications of Computer Vision (WACV), 2020.
6. Kolkar R & V, G. Human activity behavioural pattern recognition in smart home with long-hour data collection. *Journal of Information System and Technology Management (JISTM),* 2021.
7. Chen JIZ. Smart security system for suspicious activity detection in volatile areas. *Journal of Information Technology and Digital World (JITDW),* 2020.
8. Ketenci G, Kurt T, Önal S & Erbil C. Suspicious activity detection for anti-money laundering. IEEE, 2021.
9. Genemo MD. Suspicious activity recognition for monitoring cheating in exams. Proceedings of the Indian National Science Academy, 2022.
10. Sakib S, Rahman A, Haque F, Hossain SB & Masud M. Unusual driver behavior detection in videos using deep learning models. *International Journal of Management, Technology, and Social Sciences (IJMTS),* 2022.
11. Ullah I, Hussain M, Aboalsamh H & Ullah A. Anomaly detection in surveillance videos using deep learning. IEEE Access. 2022; 10:65832–65845.
12. Sultana M, Sufian A & Dutta P. Real-time abnormal activity detection in video surveillance using convolutional neural networks. *Journal of Visual Communication and Image Representation*. 2021; 80:103279.
13. Singh S, Saini M & Yadav SK. A hybrid deep learning approach for suspicious activity detection in crowded environments. *Multimedia Tools and Applications.* 2023; 82(3):3567–3590.
14. Zhang Y, Chen Z & Wang X. SlowFast network-based suspicious activity recognition for intelligent surveillance systems. *International Conference on Artificial Intelligence and Soft Computing (ICAISC),* 2023.