Real-Time Shoplifting Detection

Rahul
School of Computer Science and Engineering
Galgotias University,
Greater Noida, Uttar Pradesh- 201310
Email: rahul.21scse1011025@galgotiasuniversity.edu.in

Vikas Singh
School of Computer Science and Engineering
Galgotias University,
Greater Noida, Uttar Pradesh- 201310
Email: vikas.21scse1011026@galgotiasuniversity.edu.in

Abstract: Shoplifting poses a substantial financial challenge for retailers globally. Traditional surveillance methods lack the real-time capability needed for effective intervention, often requiring extensive manual monitoring. In response, this paper presents an automated solution utilizing artificial intelligence (AI) and machine learning (ML) techniques, developed with a hybrid neural network and deployed through a Flask-based backend. The neural network combines convolutional layers for spatial feature extraction with gated recurrent units (GRU) for temporal sequence analysis, classifying video segments as shoplifting or non-shoplifting events. Flask enables seamless deployment, offering a real-time interface for practical implementation in retail environments. Tested on the UCF-Crime dataset, the model achieved 93% accuracy, demonstrating its efficacy and potential for deployment in retail environments.

1. Introduction:

Retail theft, particularly shoplifting, continues to cost businesses billions annually. Existing video surveillance systems rely heavily on manual monitoring, limiting their effectiveness in real-time crime prevention. AI-driven video analysis, capable of detecting suspicious behavior, presents a promising alternative. Convolutional Neural Networks (CNNs) have revolutionized object detection, making them suitable for identifying behaviors in image sequences. However, for video surveillance where temporal data is crucial, Recurrent Neural Networks (RNNs), particularly Gated Recurrent Units (GRUs), can enhance the model's understanding of events unfolding over time.

This paper introduces an AI-powered solution to detect shoplifting by combining CNN and GRU architectures for video classification The Flask framework is utilized as the backend enabling real-time interaction with the system for efficient deployment in retail. Our aim is to develop a system that can process live video feeds, identify potential shoplifting incidents, and alert staff immediately, reducing financial losses and enhancing security operations.

2. Literature Survey:

Shoplifting detection via machine learning in video surveillance is a rapidly growing field, with CNN and RNN models showing significant promise in prior studies.

(i) **CNN for Spatial Analysis**:
Convolutional Neural Networks (CNNs) have been widely applied to detect objects and classify actions in static images, excelling in spatial feature extraction.

Research shows that CNNs are effective in isolating important visual elements from video frames, such as human body movements and object handling, which are essential for identifying shoplifting behaviors.

(ii) RNNs and GRUs for Temporal Analysis: Recurrent Neural Networks (RNNs), specifically GRUs and LSTMs, have proven effective in processing temporal sequences, making them suitable for analyzing behavioral patterns over time. GRUs are often preferred for their computational efficiency while handling long-term dependencies in video sequences. Studies have demonstrated that hybrid models combining CNN and RNN layers achieve higher accuracy for action recognition in videos.

(iii) **Flask as a Real-Time Deployment Framework**: Flask, a lightweight Python web framework, has gained popularity in deploying ML models for real-time applications. Its RESTful API capabilities allow seamless integration with front-end applications, making it ideal for surveillance systems where instant alerts and data streaming are essential.

This study builds on existing knowledge by integrating CNN-GRU architectures with Flask to develop a real-time, deployable solution for shoplifting detection.

3. Problem Statement: The major challenge in shoplifting detection is the real-time processing and classification of video data to identify suspicious behavior without extensive manual effort. This paper addresses this by developing an automated model capable of real-time classification, reducing the need for manual intervention and minimizing theft-related losses.

4. Methodology : This section outlines the approach used to develop the model and backend for the shoplifting detection system.

4.1 Data Collection : The UCF-Crime dataset, a well-known dataset for anomaly detection in surveillance footage, serves as the primary data source. This dataset includes over 1900 videos covering various crime categories, including shoplifting. The videos are pre-labeled as "shoplifting" or "non-shoplifting," providing a foundational dataset for model training. To increase sample diversity, each video is segmented into smaller sequences, yielding a larger set of labeled examples for training and testing.

4.2 Model Development

Architecture: The model combines a CNN for spatial feature extraction and a GRU for temporal sequence analysis. The CNN processes each frame to extract spatial features, such as movement patterns and object interactions, which are crucial in identifying shoplifting. The GRU then processes the sequential output from the CNN to classify each video segment as "shoplifting" or "non-shoplifting."

Transfer Learning: To improve accuracy and reduce training time, the CNN is initialized with pretrained weights from ImageNet, enabling the model to leverage knowledge from a vast dataset. The GRU processes these features to analyze patterns across frames.

Training: The model is trained using TensorFlow, with binary cross-entropy as the loss function and accuracy as the primary evaluation metric. Several configurations were tested to optimize the model, with hyperparameters such as batch size, learning rate, and the number of GRU units adjusted for optimal performance.

4.3 Backend Development

Flask Integration: Flask serves as the backend, connecting the trained model to a real-time video input interface. Flask's lightweight structure and RESTful API support allow easy communication between the model and front-end application, enabling real-time alerting and data handling.

API Design: The Flask API endpoints are structured to handle video frame data and return predictions. An endpoint accepts video input, processes it through the model, and provides instant feedback, allowing retail staff to respond quickly to detected shoplifting incidents.

Deployment: The Flask application is deployed on a cloud server for scalability, allowing multiple stores to access the model simultaneously without compromising response time.

5. Experiments

Several experiments were conducted to test the model's accuracy, precision, and latency in a simulated retail environment.

Hyperparameter Tuning: The model underwent tuning of parameters such as batch size, learning rate, and frame sequence length to achieve optimal performance.

Validation and Testing: The dataset was split into training (70%) and testing (30%) subsets, with additional validation during training. The Flask application was tested for latency, ensuring that predictions could be generated within a 1-2 second window.

6. Results and Discussion

Comparison with Existing Methods: The achieved accuracy surpasses traditional shoplifting detection systems, which often rely on simpler motion detection algorithms with limited predictive accuracy.

Flask Performance: The Flask-based backend maintained a low latency, allowing near-instantaneous responses, critical for practical deployment in busy retail environments.

7. Conclusion:

This research demonstrates the feasibility of using AI/ML for real-time shoplifting detection. By combining CNN-GRU architectures with a Flask backend, the model efficiently identifies suspicious behaviors, allowing store staff to respond to incidents in real-time. Future improvements may focus on increasing model resilience to varied lighting and camera angles, improving generalizability across different store layouts.

References:

- [1] Kirichenko, L., et al. “Detection of Shoplifting on Video Using a Hybrid Network.” *Computation*, 2022.
- [2] Rehman, A., Belhaouari, S.B. “Deep Learning for Video Classification: A Review.” *TechRxiv*, 2021.
- [3] Gollapudi, S. *Learn Computer Vision Using OpenCV: With Deep Learning CNNs and RNNs*. Apress, 2019.
- [4] Yamato, Y., Fukumoto, Y. “Security Camera Movie and ERP Data Matching System to Prevent Theft.” *Consumer Communications & Networking Conference*, IEEE, 2017.