**Real-time Anomaly Detection in Video Surveillance**

Bao Ngo

School of Science and Engineering

University of Missouri-Kansas City

bndbv@mail.umkc.edu

Feng Zheng

School of Science and Engineering

University of Missouri-Kansas City

fz4dt@umsystem.edu

## Introduction

The increasing use of surveillance cameras in the United States, driven by technology and security concerns, has led to a reliance on human operators for monitoring and anomaly detection. However, human operators have limitations such as fatigue and distraction. To address this, real-time automatic anomaly detection in video surveillance is crucial. This system can continuously monitor and promptly respond to unusual activities, enhancing security by detecting crimes like shootings, shoplifting, assault, and robbery. Furthermore, it serves as an early threat detection tool, identifying suspicious events like robbery, unusual vehicle movements, etc. Rapid anomaly detection enables quicker responses from security personnel and law enforcement, reducing the risk of danger or damage.

1. **Objectives**
   * Early Threat Detection: to build a real-time anomaly detection system that can identify unusual events and quickly respond to possible threats and reduce the chance of damage.
   * Scalability: to build a real-time anomaly detection system that can process multiple video surveillance streams at once, which provides a scalable solution that retains accuracy.
   * Reduced False Positives: to build a real-time anomaly detection system that can learn from past data and normal behavior patterns to decrease the number of false positives.
   * Emergency Notification: to build a real-time anomaly detection system that can send a notification to users when an anomaly action is detected.

## Data Preprocessing

Preprocessing video data for real-time anomaly detection in video surveillance involves several steps to ensure that the data is prepared, cleaned, and transformed in a way that facilitates accurate anomaly detection. Here are the key steps:

* + **Data Collection:** Gather video data from surveillance cameras, which can be in the form of video streams or recorded footage. Ensure that the data is timestamped to maintain temporal information.

A screenshot of a computer

Description automatically generated

Figure 1: Data Collection

* + **Frame Extraction:** To work with video streams, extract individual frames from the video to process them individually.
  + **Removing Duplicate Frames:** Eliminate identical frames to reduce redundancy. This will significantly decrease the number of frames to process. We used the absolute different between the frames to achieve this.

Several images of people sitting at a desk

Description automatically generated

Figure 2: Absolute difference between frames

* + **Refactored Filtered Video:** After removing duplicate frames, we refactored the video used the filtered frames. This will effectively reduce the amount of work on prediction stage.

A screen shot of a computer code

Description automatically generated

Figure 3: Refactored video using frames

* + **Data Normalization:** Normalize the extracted features to have consistent scales and distributions. This step is crucial for ensuring that different features are treated equally during anomaly detection.
  + **Data Splitting:** Splitting the data into training, and testing datasets. So that we can train the model with these datasets.

A screenshot of a black screen

Description automatically generated

Figure 4: Splitting data with txt file

## Data Analysis

* + **Temporal Analysis:** Analyze the temporal behavior of objects or regions of interest across multiple frames. Detect anomalies by identifying deviations from typical object behavior over time.
  + **Thresholding:** Set appropriate thresholds for anomaly scores or probabilities to classify detected anomalies as genuine or false positives.
  + **Model Updating:** Periodically update the background models, feature extraction methods, and anomaly detection algorithms to adapt to changing environmental conditions and maintain accuracy.
  + **Evaluation and Fine-tuning:** Continuously evaluate the performance of the anomaly detection system and fine-tune parameters and algorithms to minimize false positives and improve detection accuracy.

## Notification and Logging

* + **Alert Generation:** Generate alerts or notifications when anomalies are detected. This may involve sending alerts to security personnel or triggering other actions to address the anomaly. Some ways are through email, message through subscripted phone, and notification through application.
  + **Logging and Storage:** Log the detected anomalies and associated information for future analysis and auditing. Store the preprocessed video data, anomalies, and alerts for reference and evidence.

## Methodology

* **I3D (Inflated 3D ConvNet)**
  + 3D Convolutional Layers: I3D uses 3D convolutional layers that operate on video data, capturing spatial and temporal features simultaneously.
  + Inflated Weights: Initializing the 3D convolutional layers with 2D convolutional filters pretrained on large-scale image classification datasets. These pretrained weights helps in transfer learning and faster convergence.
  + Two-Stream Architecture: I3D typically employs two parallel streams of networks: one for RGB frames and another for optical flow. This enables the model to analyze both frame content and motion patterns.

A diagram of a computer network

Description automatically generated

Figure 5: The training process for the two-stream I3D on Kinetics Dataset. Image by author, adapted from Carreira and Zisserman (2017) [2].

* **LIME**
  + Model Agnostic: LIME is model-agnostic, meaning it can be used with any machine learning model, whether it's a decision tree, neural network, random forest, or any other model. It does not rely on the model's internal structure.
  + Local Interpretability: LIME focuses on explaining the prediction of a single instance or a small group of instances. It doesn't provide global explanations for the entire model.
  + Human-Readable Explanation: The local model's coefficients or rules can be interpreted by humans, providing insights into why the model made a particular prediction for a given instance.

A collage of different dogs

Description automatically generated

Figure 6: Example of the Lime model

## Proposed Approach

## A diagram of a computer network Description automatically generated

Figure 7: The flow diagram of proposed anomaly detection [1]

## In Figure 7, we change C3D model to I3D model to extract features for each video segment. In addition, for the feature extraction part, we are using the pre-calculated I3D and then we build train a fully connected neural network to calculate the ranking loss between the highest score instances in the positive bag (containing anomaly) and the negative bag (containing no anomaly)

## Model Conception

A screen shot of a computer program

Description automatically generated

## Figure 8: Custom Dataset class

## In Figure 8, we create the custom dataset class to load the UCF dataset we have. This class contain parameter to determine if the dataset was normal or anomaly and also the train test status.

1. **Model**

## A screen shot of a computer program Description automatically generated

## Figure 9: Fully connected neural network to compute the ranking loss

## In Figure 9, we create a neural network model using the neural network. Module base class. This model consists of several fully connected layers (linear layers) along with activation functions and dropout layers.

## A computer screen shot of a program code Description automatically generated

## Figure 10: Training function

## In Figure 10, we create a training loop for a neural network model. We use two data loaders, nloader and aloader to train the model.

## Conclusion

In conclusion, this paper has underscored the critical importance of real-time anomaly detection in video surveillance systems, emphasizing its role in enhancing societal security. We have discussed the significance of early threat detection, scalability, reduced false positives, proactive response, and emergency notification as key objectives in the context of this technology.

## References

[1] Sultani, Waqas, et al. “Real-World Anomaly Detection in Surveillance Videos.” CRCV, www.crcv.ucf.edu/projects/real-world/. Accessed 6 Dec. 2023.

[2] Carreira, J., & Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset. In proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 6299–6308)