October 2023

Anomaly Detection in Compressed Video Streams Using Machine Learning

Yuriy Aleksandrovich Romanenko

Anomaly Detection in Compressed Video Streams Using Machine Learning

ABSTRACT

Video outlier or anomaly detection is currently performed using computer vision and image processing techniques such as filtering, motion estimation, etc. that require decompressed video streams. Applying machine learning techniques directly to a video stream is computationally expensive. This disclosure describes the use of a pre-trained machine learning model that can analyze compressed video bitstreams directly without decompression to identify likely anomalies in the video. Such direct analysis of compressed video streams can be performed at significantly lower computational cost while still yielding a strong anomaly detection signal. Sequences in which anomalies are detected can be decompressed and further analyzed using conventional techniques of video analysis. Sequences over longer time spans (and associated groundtruth information about whether an anomaly exists in the sequence) can be used for continuous unsupervised training of the same model. This can improve detection of anomalies in subsequent sequences to enable better anomaly detection, without requiring collection of additional training data or separate training cycles.

KEYWORDS

- Anomaly detection
- Video anomaly
- Compressed bitstream
- Surveillance camera
- Security camera
- Unsupervised learning
- Anomaly detection

BACKGROUND

Video outlier or anomaly detection is important in many applications. For example, analysis of video from a surveillance camera or other video recording devices can reveal intrusion or other actions that are unusual (deviate from normal patterns). Currently, video anomaly detection relies on computer vision and image processing techniques such as filtering, motion estimation, etc. applied to decompressed video streams. Applying machine learning techniques directly to a video stream is computationally expensive.

DESCRIPTION

Video streams produced by a camera (e.g., surveillance camera or other cameras) may be transmitted as a video stream to a digital video recorder (DVR) or network video recorder (NVR) for storage. Such streams are usually compressed using video codecs that utilize image-to-image prediction, motion estimation, etc. This disclosure describes machine learning techniques that can analyze the compressed bitstreams directly without decompression to identify likely anomalies in the video. Such direct analysis of compressed video streams can be performed at significantly lower computational cost while still yielding a strong anomaly detection signal.

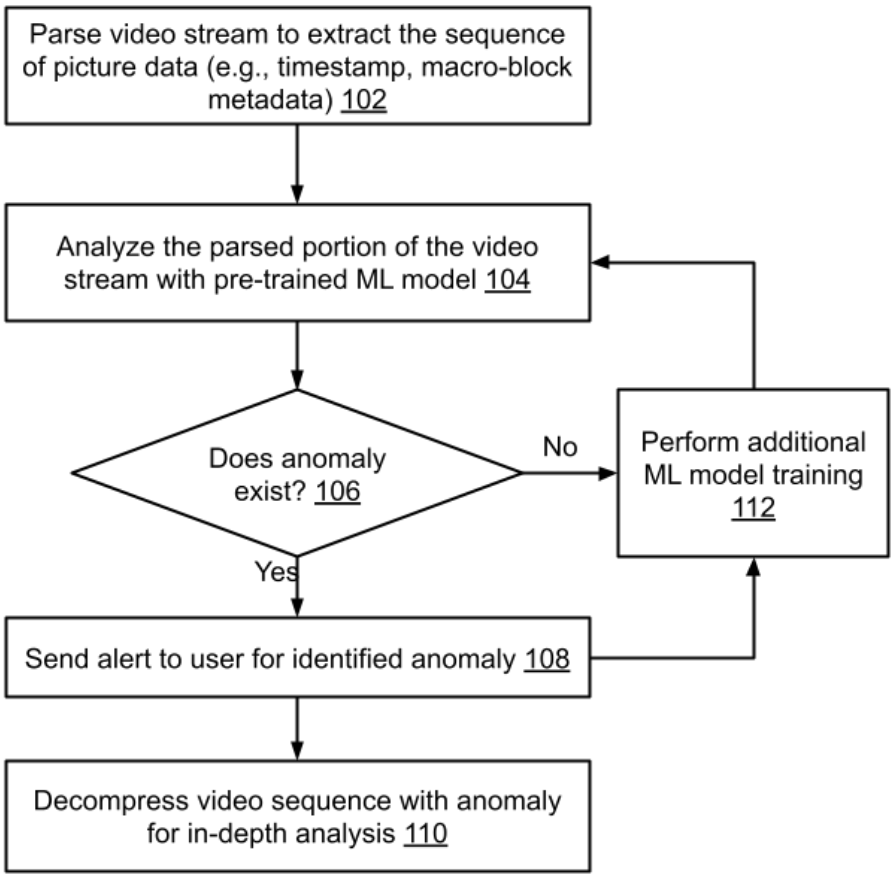

Fig. 1: Anomaly detection in compressed video stream using machine learning

Fig. 1 illustrates a method to identify anomalies in compressed video streams using a machine learning model. A compressed video bitstream is parsed but not decompressed (102). Sequences of compressed picture data are extracted from the parsed stream. These can include, e.g., slice and macro-block metadata, timestamp, motion vectors, residual code volume, etc. The extracted data is fed to a pre-trained machine learning model (104). The machine learning model is trained on video streams that lack anomalies.

Based on the extracted data, the model outputs (106) an indication of whether an anomaly likely exists in the sequences of the video stream. If an anomaly is detected, an alert can be raised (108). For example, if the video stream is from a security system, a motion or

intruder alert may be raised. Alternatively, or in addition, the particular sequence can be decompressed (110) and analyzed in-depth, e.g., using any suitable video analysis techniques to determine if the anomaly is indeed present. The in-depth analysis is thus limited only to portions where analysis of the compressed bitstream of the video indicates an anomaly, which is computationally more efficient than decompressing the entire bitstream.

Optionally, sequences over longer time spans (and associated groundtruth information about whether an anomaly exists in the sequence) can be used for continuous unsupervised training (112) of the same model. This can improve detection of anomalies in subsequent sequences to enable better anomaly detection, without requiring collection of additional training data or separate training cycles. The described techniques can be used in security cameras, network video recorder (NVR) devices, etc.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (e.g., information about a video feed from a surveillance camera or other device, or video recordings), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes the use of a pre-trained machine learning model that can analyze compressed video bitstreams directly without decompression to identify likely

anomalies in the video. Such direct analysis of compressed video streams can be performed at significantly lower computational cost while still yielding a strong anomaly detection signal. Sequences in which anomalies are detected can be decompressed and further analyzed using conventional techniques of video analysis. Sequences over longer time spans (and associated groundtruth information about whether an anomaly exists in the sequence) can be used for continuous unsupervised training of the same model. This can improve detection of anomalies in subsequent sequences to enable better anomaly detection, without requiring collection of additional training data or separate training cycles.

REFERENCES

1. Guo, Jianting, Peijia Zheng, and Jiwu Huang. "Efficient privacy-preserving anomaly detection and localization in bitstream video." *IEEE Transactions on Circuits and Systems for Video Technology* 30, no. 9 (2019): 3268-3281.
2. Duong, Huu-Thanh, Viet-Tuan Le, and Vinh Truong Hoang. "Deep Learning-Based Anomaly Detection in Video Surveillance: A Survey." *Sensors* 23, no. 11 (2023): 5024.