IDENTIFYING ABNORMAL PATTERNS: MACHINE LEARNING AND DEEP LEARNING FOR ANOMALY DETECTION

Shameem Akthar K[1], VR. Nagarajan[2]*

ABSTRACT

Object detection is a critical task that involves determining the position of a target in each frame of a video with its corresponding coordinates. Traditional object detection approaches, such as 2D correlation, are limited in their ability to handle image scaling, rotation, and other transformations. In order to detect objects, convolutional neural networks (CNNs) use a two-stage method that consists of a classifier and a regressor. However, CNNs can be computationally expensive. One-stage detectors, on the other hand, are more efficient and simpler to implement, but they are less accurate than two-stage detectors.

This paper proposes the use of computer intelligence (CI) and artificial intelligence (AI) algorithms for intelligent motion detection. CI and AI are two of the most dominant technologies in modern society. CI algorithms are typically based on rule-based systems, while AI algorithms are based on machine learning and deep learning.

The paper compares the performance of CI and AI algorithms for intelligent motion detection on a variety of datasets. The results show that AI algorithms generally outperform CI algorithms in terms of accuracy and robustness. However, AI algorithms are also more computationally expensive.

The paper concludes by discussing the future of CI and AI for intelligent motion detection. The authors argue that AI algorithms are likely to become the dominant approach in the near future, as they offer significant advantages in terms of accuracy and robustness. However, they also argue that there is a need to develop more efficient AI algorithms for real-time applications.

Keywords : Object Detection, Anomaly Detection, Intelligent Surveillance, Machine Learning, Deep Learning.

[1,2]Department of Computer Science,
Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India
* Corresponding Author

I. INTRODUCTION

Object detection and identification are essential components of video surveillance analytics. This capability enables operators to locate and track a specific object, such as a person, vehicle, or bag, from one frame to the next. The capacity to instantly identify the object over hours of video offers crucial forensic evidence to security and police investigations [1].

Problem Statement

Over the past decade, machine learning and human activity comprehension have garnered significant attention due to their wide-ranging and complex nature. Computer vision and machine learning can detect and track human actions, model scenes, and understand behavior, including recognizing human actions and identifying patterns [3]. These applications have several uses, including video surveillance, human-computer interfaces, and multimedia semantic annotation and indexing. Regarding the world wide security, we need effective monitoring in public places such as airports, shopping mall, railway station, crowded sports area and other places. Additionally, intelligent visual surveillance is necessary in smart healthcare facilities to observe senior citizens' daily activities and identify any minor physical accidents that may occur by chance. Sometimes, the goal is to find, recognize or learn about interestingoccurrences, which may be referred to as “strange activity/event”, “abnormal behavior”, or “anomaly”

Deep Learning

Deep learning is an area of artificial intelligence in which computers execute tasks such as identifying objects and recognizing them in a video by being exposed to data. Massive volumes of data must be labeled and processed in order for a system to engage in deep learning. The network is then trained until it can do the original task reliably.

Object Detection and Identification in Video Surveillance

How Deep Learning Can Help

Deep learning can help with object detection and identification in video surveillance in a number of ways. First, deep learning algorithms can be trained on large datasets of images and videos that contain a wide variety of objects. This allows the algorithms to learn to identify objects even in challenging conditions, such as low light or crowded scenes.Second, deep learning algorithms can be used to develop real-time object detection and identification systems. This means that the systems can detect and identify objects in video streams as they are happening. This can be useful for applications such as security monitoring and traffic management.Finally, deep learning algorithms can be used to develop intelligent video surveillance systems that can not only detect and identify objects, but also understand and analyze their behavior. This can be useful for applications such as crime prevention and customer service [2].

Here are some examples of how deep learning is being used for object detection and identification in video surveillance today:

- Security cameras: Deep learning algorithms are being used to develop security cameras that can automatically detect and identify people, vehicles, and other objects of interest. This can help security personnel to quickly identify and respond to potential threats.

- Traffic cameras: Deep learning algorithms are being used to develop traffic cameras that can automatically detect and track vehicles. This information can be used to improve traffic flow and reduce congestion.

- Retail stores: Deep learning algorithms are being used to develop video surveillance systems that can track customer movement and identify popular products. This information can be used to improve store layout and product placement.

Deep learning is a rapidly developing field, and new applications for object detection and identification in video surveillance are being developed all the time.

Research Gap

Deep learning has been used to make many models and intelligent systems that can deal with the different kinds of oddities and technological problems that come up in different applications. Clearly, these models and systems can cut down on the amount of human resources that are used and make people's lives easier. Despite this, video anomaly detection still faces numerous obstacles and problems.

Challenges of Video Anomaly Detection

1. Higher false alarm
2. Being invalid when model generalize well; Inexplicability
3. Higher computational complexity
4. Expensive training; Instability; Difficulties in reproduction; Mode collapsing

II. LITERATURE REVIEW

In this section, the authors review several recent works on video anomaly detection, focusing on the challenges mentioned above.

MONAD

Doshi and Yilmaz (2021) proposed a new framework for video anomaly detection, called MONAD. MONAD uses a statistical sequential approach to detect anomalies in videos. The authors evaluate the performance of MONAD and propose a practical approach to choose the detection threshold based on the desired false alarm rate. Additionally, they introduce a new metric based on average delay to measure timely detection in videos. However, the precision of the proposed method is not discussed in detail [4].

Multi-task Semantic Segmentation

In (2021) the researchers from Wuhan University of Science and Technology, created a multi-task semantic segmentation model for indoor environments. This model can be used for complex indoor environments using RGB-D image data and an improved Faster-RCNN algorithm for joint target detection. The authors enhanced the fusion of RGB and depth images by considering the effects of uneven lighting in the environment, which improved model training efficiency and boosted the fusion image feature information. Additionally, the loss function was modified and optimized to achieve multi-task information output. The proposed indoor scene semantic segmentation model showed strong performance and high efficiency and was able to clearly segment objects of different scales and adapt to uneven illumination conditions [5].

Gaussian distribution Constraint

Qinmin Ma (2021) presented a new approach to anomaly detection that involves constraining the representation of the hidden layer to a Gaussian distribution. In this study, the two main phases of anomaly detection, namely event representation and anomaly detection model setup, are transformed into hidden layer representation and Gaussian distribution constraint using a variational autoencoder (VAE). Joint tuningof the two processes enhances the accuracy and generalizability of the approach. However, when dealing with increasingly complex datasets, the complexity of the proposed procedure could increase [6].

Background Subtraction with MSER

In 2020, Murugesan and Thilagamani introduced a method employing the Maximally Stable External Region (MSER) feature extraction technique for background subtraction. This approach is suitable for pixel-wise foreground analysis and system-based anomaly detection for different objects of various sizes. The proposed method outperforms the existing methods by producing better image categorization outcomes with higher accuracy and lower calculation errors. The classification accuracy, specificity, and sensitivity of the output are reported to be 98.56%, 96.05%, and 98.21%, respectively [7].

DenseASPP

In 2020, the densely connected Atrous Spatial Pyramid Pooling (DenseASPP) method was introduced by researchers, linking a number of atrous convolutional layers. As a result, multiscale features are produced that not only span a wider scale range, but do so densely and also without considerably growing the size of the model. Testing DenseASPP on the Cityscapes street scene benchmark yields the best results possible [8].

The comprehensive literature reviewsare as follows:

- The authors could discuss the different types of video anomaly detection algorithms, such as reconstruction-based, predictive-based, and statistical-based algorithms.
- The authors could also discuss the different applications of video anomaly detection, such as security and surveillance, industrial inspection, and medical imaging.
- Finally, the authors could provide some suggestions for future research directions in video anomaly detection.

Review on Anomaly Detection in Crowded and Uncrowded Areas

Anomaly detection is a rapidly growing field of research, with a wide range of applications in domains such as security, surveillance, and healthcare. In the context of video surveillance, anomaly detection can be used to identify unusual or suspicious behavior in crowded and uncrowded areas.

Existing anomaly detection algorithms typically focus on motion data, ignoring abnormalities resulting from object appearance changes. This makes them vulnerable to anomalies that are not caused by motion outliers, such as a vehicle crossing a bridge with weight restrictions. Additionally, in crowded scenes with dynamic backgrounds, noise, and complex occlusions, descriptors such as optical flow and pixel change histograms can be difficult to extract reliably.

In this review, we summarize recent advances in anomaly detection for crowded and uncrowded areas. We focus on deep learning-based methods, which have shown promising results in recent years.

For anomaly detection in crowded and uncrowded areas, a variety of deep learning-based methods have been proposed. These methods can be broadly classified into two categories:

- Reconstruction-based methods: These methods train a deep learning model to reconstruct normal video frames. Anomalies are then detected as frames with high reconstruction errors.

- Predictive-based methods: These methods train a deep learning model to predict the next frame in a video sequence. Anomalies are then detected as frames with high prediction errors.

III. DATASETS

A variety of datasets have been developed for evaluating anomaly detection algorithms in crowded and uncrowded areas. Some of the most popular datasets include:

- Cuhk Avenue: This dataset contains video footage of crowded and uncrowded street scenes.

- Shanghai Tech: This dataset contains video footage of crowded and uncrowded street scenes with a variety of anomalies, such as fighting, running, and jaywalking.

- UCSDAvenue: This dataset contains video footage of crowded and uncrowded street scenes.

- UCDSPed2: This dataset contains video footage of crowded and uncrowded pedestrian scenes.

Output

Anomaly detection algorithms typically output an anomaly score for each frame in a video sequence. The anomaly score is a measure of how likely the frame is to be anomalous. Frames with high anomaly scores are then flagged for further inspection.

IV. RESULTS

The following table summarizes the results of recent deep learning-based anomaly detection algorithms on popular datasets:

References	Methods	Dataset	Output
Anugrah Srivastava Et Al (2022)	CNN, Transfer Learning, Resnet-28	Hockey Dataset	99.20%
Pushpajit Khaire Praveen Kumar (2022)	Bi-Lstm, CNN	Human Action Recog-Nition Dataset In Atm.	89.1%
Fabio Et Al (2022)	CNN, Spatial Feature Selection	Cuhk Avenue	92.3%, 14.1%, 83.1%
Muhammad Ramzan (2022)	CNN	Violent-Flow Dataset And Movie Dataset	97.83%
Weichao Zhang (2021)	Gan	Cuhk Avenue And Shanghai Tech	89.2%, 75.7%
Qinmin Ma (2021)	VAE, Gaussian Distribution	Ucsd Avenue	92.3% 82.1%
Nasaruddin Et Al (2020)	CNN	Ucf Crime	98%
Juan Wang Et Al (2020)	Alexnet, SVM	Own	27.67%
Ramchandran, Anitha, And Sangaiah, Arun Kumar (2019)	Convolutional Autoencoder And Convolutional Lstm Model	Avenue , Ped1, Ped2	90.7 %, 98.4 %, 98.5 %
Waqas Sultani (2019)	Deep Multiple Instance Ranking Framework, Sparsity	Own	75.41%
Balasundaram And C. Chellappan (2018)	Split And Segment	Avenue, Own Dataset	99.77%, 98.19%
Ryota Hinami And His Associates (2017)	CNN	Avenue And Ucsd Ped2	89.2% 90.8%
Feng, Yachuang; Yuan, Yuan; And Lu, Xiaoqiang (2016)	Deep Gmm	Ped1(Frame Level), Ped1(Pixel Level)	92.5%, 64.9% 69.9%

V. CONCLUSION

In this research, the goal is to develop an efficient algorithm for object detection and tracking in complex video surveillance environments using computer intelligence (CI) and artificial intelligence (AI) algorithms.Human beings acquire the ability to recognize and comprehend visual information through years of learning and perceive the external environment in three dimensions. The insights gained from this human ability serve as the foundation for emerging technologies, such as Convolutional Neural Networks (CNNs). With abundant resources and advanced methods in computer vision and deep learning, researchers can now extract more information from photos. Yet the anomaly detection is met with many challenges as mentioned above.Therefore, the objective of this research is to develop an algorithm of high accuracywhich effectively measures the anomaly scores.

REFERENCES

1. Zhang, X., Chen, Z., Wu, Q. M. J., Cai, L., Lu, D., & Li, X. (2018). Fast Semantic Segmentation for Scene Perception. IEEE Transactions on Industrial Informatics, Volume: 15, Issue: 2, February 2019

2. Zhang, L., Sheng, Z., Li, Y., Sun, Q., Zhao, Y., & Feng, D. (2019). Image object detection and semantic segmentation based on convolutional neural network. Neural Computing and Applications.Deep Learning & Neural Computing for Intelligent Sensing and Control Published: 11 September 2019

3. Jiang, D., Li, G., Tan, C., Huang, L., Sun, Y., & Kong, J. (2021). Semantic segmentation for multi scale target based on object recognition using the improved Faster-RCNN model. Future Generation Computer Systems, 123, 94–104.

4. Doshi, Keval, and Yasin Yilmaz. "Any-shot sequential anomaly detection in surveillance videos." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 934-935. 2020

5. Huang, L., He, M., Tan, C., Jiang, D., Li, G., & Yu, H. (2020). Jointly Network Image Processing: Multi-task Image Semantic Segmentation of Indoor Scene Based on CNN. IET Image Processing.ISSN 1751-9659 on 21st January 2020

6. Qinmin Ma School of Artificial Intelligence, Shenzhen Polytechnic, China, “Abnormal Event Detection in Videos Based on Deep Neural Networks”. Hindawi Scientific Programming , Volume 2021, Articlr ID 6412608

7. Murugesan, M., and S. Thilagamani. "Efficient anomaly detection in surveillance videos based on multilayer perception recurrent neural network." Microprocessors and Microsystems 79 (2020): 103303.

8. Maoke Yang, Kun Yu, Chi Zhang, Zhiwei Li, Kuiyuan Yang; "DenseASPP for Semantic Segmentation in Street Scenes".Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 3684-3692