Multiple Object Tracking

Rittin Mithra C, Sugash Srimari R, Sabarishwaran S

UG Scholar Amrita School of Computing, Amrita Vishwa Vidyapeetham – Chennai

Abstract :- **In this paper we are tracking the multiple objects in a video where motion patterns are represented through bag-of-visual-words constructed from optical flow histograms. Anomalies are detected at multiple intervals, frame-level as well as pixel-level. The proposed approach is evaluated on the benchmark UCSD Ped1 dataset, achieving promising results. A comprehensive evaluation using precision, recall, and F1 score is done. We propose a novel model to track multiple objects in the range of interest in surveillance, fraud monitoring and disaster management videos. An experimental evaluation is conducted with a new dataset of crowded scenes, composed of 36 video sequences and five well defined abnormality categories. The proposed representation is shown to outperform various state of the art Multiple Object Tracking techniques.**

Keywords : OpenCV, MOG, USCD Ped1 Dataset, Decision Tree Classification.

I.INTRODUCTION

Multiple Object Tracking is important for surveillance, error detection, fraud monitoring, and disaster management applications. Most existing video object tracking techniques detect anomalies at the frame level and ignore temporal relationships between events. To localize anomalous events in videos, we propose a Multiple objects tracking framework that considers both spatial and temporal cues. We formulate Multiple objects tracking as a classification problem and train a decision tree classifier using handcrafted pixel-level features. Features encode motion, foreground, spatial, and statistical properties of video cells that can distinguish between normal and abnormal regions.

. For detection, features are extracted from the current and previous frames and classified to obtain an anomaly score for each cell. Cells with values above the detection threshold are flagged as anomalous regions in the output highlight. We evaluate the proposed method using the UCSD Ped1 dataset, which contains videos of pedestrians with occasional anomalous events such as running, loitering. We propose a spatiotemporal framework for and demonstrate its effectiveness for pedestrian videos. This approach can detect a variety of anomalous events in surveillance, fraud monitoring, fault diagnosis, and disaster management video.

II.LITERATURE REVIEW & PROBLEM DESCRIPTION

The problem of multiple object tracking has been studied extensively and numerous solutions/approaches have already been proposed. {*A-D*} involves the previous methods proposed and the drawbacks.

A .Anomaly detection in crowded scenes

In this paper it solves the problems of many anomalies detection and it concentrate on the issue of representation, namely, how to design localized video representations that enable anomaly detection in crowded scenes. This precludes any form of global statistical inference, using MRFs, LDA, or any such models: while these can certainly improve performance, they tend to mask the limitations of the underlying visual representation. We identify three properties that the representation must have:

1) jointly model appearance and dynamics of crowd patterns
2) ability to detect temporal and
3) spatial abnormalities [1]

B. Multiple object tracking method using Kalman filter

In this work, the researchers established an object motion model utilizing the Kalman filter, which reduces the search area and search time of moving objects to accomplish quick tracking by predicting the position of the item based on its current information. Creating a comparable connection by shifting object characteristics together to address separation after objects have merged. The technique of Kalman filtering is used for locating objects in films. It is a recursive method that makes use of noisy data to estimate an object's state over time. To forecast an object's position, the Kalman filter models the object's motion and considers previous information about the object's behaviour. [2]

C. Multiple Object Tracking Using Particle Filters

In this work, the researchers established an object motion model utilizing the Particle filter, they can handle nonlinear and non-Gaussian distributions. The basic idea behind particle filters is to represent the distribution of the state of each object with a set of particles, where each particle corresponds to a possible state of the object. The particles are propagated forward in time, and their weights are updated based on how well they match the observations.The drawback of using this filter is computational complexity[3]

.

D. Video anomaly detection and localisation based on the sparsity and reconstruction error of auto-encoder

A fast and accurate video anomaly detection and localisation method is presented. Two novel cubic-patch-based anomaly detectors are introduced, one based on the power of an auto-encoder (AE) and one based on the power of sparse representation of an input video patch. The experiment results show that the method has a better performance than state-of-the-art methods on two UMN and UCSD benchmarks. [4]

E. Multi-Object Detection and Tracking, Based on DNN, for Autonomous Vehicles: A Review

In this study it evaluates the state of-the-art techniques that address this challenge, with three primary sensors camera, LiDAR, and RADAR with DNN, and fusion of sensor data with DNN. The analysis shows that there exists an excellent potential to design a more optimized solution to address this challenge. a graph structure where it maintains multiple hypotheses about the number and trajectories of the objects in the video. The image information drives the process of e his work proposes a perception model for autonomous vehicles. [5]

F. Deep Neural Networks for Object Detection

In this study we go one step further and address the problem of object detection using DNNs, that is not only classifying but also precisely localizing objects of various classes. We present a simple and yet powerful formulation of object detection as a regression problem to object bounding box masks. We dextending and pruning the graph, and the tracking process confirms and validates the detection through time. The multiple objects tracking method gives feedback, which predicts object locations to the object detection module. The most possible hypothesis provides the multiple objects tracking result. [6]

G .Multiple Object Recognition Using OpenCV

This study presents object characteristics analysis using image processing techniques. Image content characterization and supervised classifier type neural network are used in the proposed deciding method. An image data is rearranged, and a region of interest is selected during preparation. Color and texture features are extracted from an input and the device is trained to automatically identify test images. The tangent sigmoid function is used as the kernel function and the simulated results show that the used network classifier has a low error rate during training and higher classification accuracy. [7]

H .A Survey: On Multiple Object Detection and Tracking

This paper reviews research papers for object detection and tracking methods. It divides the tracking methods into three categories based on the use of object representations: point correspondence, primitive geometric models, and contour evolution. All classes require object detection at some point, with point trackers requiring detection in every frame and geometric region or contours-based trackers only when the object first appears.[8]

I . Floor Fields for Tracking in High Density Crowd Scenes

This paper presents an algorithm for tracking individual targets in high density crowd scenes containing hundreds of people. Tracking in such a scene is extremely challenging due to the small number of pixels on the target, appearance ambiguity resulting from the dense packing, and severe inter-object occlusions. The novel tracking algorithm, which is outlined in this paper, will overcome these challenges using a scene structure-based force model. [9]

III.METHODOLOGY

Decision tree classifier has shown great promise in solving the problem of MOT. In "Object Tracking" it involves Object Association and Track Management, here the detected objects are associated to frames and then the resulting tracks should be managed. Finally, the Performance of the system needs to be evaluated using standard metrics such as Precision, Accuracy, F1 Score, MOTA (Accuracy) and MOTP (Precision).

3.1 Kalman filtering:

Kalman filtering is a widely used method for tracking objects in videos. It is a recursive algorithm that estimates the state of an object over time using noisy measurements. The Kalman filter models the object's motion and incorporates prior knowledge about the object's behavior to predict its future location. However, the main drawback is it involves several complex mathematical operations, such as matrix multiplication and inversion, which can be computationally intensive and time-consuming.

3.2 Particle filtering:

Particle filtering is another popular technique for multiple object tracking. It is a Bayesian filtering method that uses a set of weighted particles to represent the object's possible state. The particles are updated iteratively based on the likelihood of the observations. Some of the disadvantages using this method are Computational Complexity as particle filter requires many particles to be sampled and propagated through the state space, which can be computationally expensive and time-consuming, particularly when tracking multiple objects in a cluttered environment.

3.3 Hybrid Methods:

Hybrid Method is another existing technique for multiple object tracking. It uses multiple tracking algorithms and fuse their results in a way that complements each other. Some disadvantages of using this method are computationally expensive, slow processing times, and selecting the combination of algorithms is a challenging task.

3.3.1 IMPLEMENTATION WORK

The main and the most important portion of this model is the use image processing techniques which are implemented, the algorithm is broadly divided into following parts:

- Data Collection
- Pre-processing
- Background Subtraction
- Track Management
- Training and Object Tracking
- Novelty Detection Techniques Measurements
- Result

3.3.1.1 Data Collection:

The first step is to collect a dataset of video sequences with multiple objects to train the neural network such as online, offline video sequences from recordings and live streams. In this project pre-captured videos (offline streams) will be taken. The images will be stored in color MP4 format.

3.3.1.2 Pre-processing:

After collection of datasets, pre-processing of the datasets is done. When a video is acquired, there may be noises present in the video. These noises affect the recognition rate greatly. So, these noises should be removed from the images. These pre-processing steps are required for the videos to be suitable for the neural network to be trained.

3.3.1.3 Background Subtraction: Using MOG2 background subtraction is done to detect objects and extract the features. Background subtraction is one of the main techniques used in object and anomaly detections.

3.3.1.4 Track Management:

After Background Subtraction using OpenCV for detecting objects for the collected datasets, Track Management is done. Associating the detected images across various frames leads to management of the resulting tracks. This includes handling occlusions, track initialization and termination, and track prediction. This is an important step as we will be only detecting objects only in the Field of View (FOV).

3.3.1.5 Training and Object Tracking:

After Track Management training the model is done. To improve the performance of the system, the neural network can be fine-tuned using the collected annotated data. This can involve transfer learning from a pre-trained object detection network or training the network from scratch. Here we initialize Object Detection by creating an object and create frames according to FPS of the dataset and IDs are assigned, IDs are added by comparing the previous and new frames.

3.3.1.6 Novelty Detection Techniques Measurements

In this section, we give an overview of the novelty detection techniques measurements used to evaluate techniques of novelty detection. In order to evaluate this, standard metrics such as detection rate, accuracy, precision, recall and F-score. Detection Rate Detection rate is a normally used measurement to check novelty detection approaches or techniques. The result of tracking can be grouped as following:

- Precision
- Recall
- F1-Score

These measurements have been obtained to check the accuracy of the model created.

IV.EXPERIMENTAL RESULTS AND ANALYSIS

The below sections provides us the results of experiments of both the techniques involved in this project.

```
Results frame wise:
Precision:  0.7910447761194029
Recall:  0.15738678544914625
F1:  0.26253869969040244
```

Figure 1: Results Obtained

A. USCD PEDESTRIAN DATA SET

In the literature, detection has frequently been evaluated by visual inspection [9], [7], [3], or with coarse ground truth, for example, frame-level annotation of abnormal events [4], [1]. This does not completely address the anomaly detection problem, where it is usually desired to localize anomalies in both space and time. To enable this, we introduce a data set1 of crowd scenes with precisely localized anomalies and metrics for the evaluation of their detection. The data set consists of video clips recorded with a stationary camera mounted at an elevation, overlooking pedestrian walkways on the UCSD campus. The crowd density in the walkways is variable, ranging from sparse to very crowded. In the normal setting, the video contains only pedestrians. Abnormal events are due to either 1) the circulation of nonpedestrian entities in the walkways, or 2) anomalous pedestrian motion patterns. Commonly occurring anomalies include bikers, skaters, small carts, and people walking across a walkway or in the surrounding grass. A few instances of wheelchairs are also recorded. All abnormalities occur naturally, i.e., they were not staged or synthesized for data set collection. The data set is organized into two subsets, corresponding to the two scenes of Fig. 5. The first, denoted “Ped1,” contains clips of 158 *238 pixels, which depict groups of people walking toward and away from the camera, and some amount of perspective distortion. The video footage of each scene is sliced into clips of 120-200 frames. A number of these (34 in Ped1) are to be used as training set for the condition of normalcy. The test set contains clips (36 for Ped1). The abnormalities of each set are summarized in Table 1.

Composition of UCSD Anomaly Data Set

Scene	Nor.	Abnormal[a]					
		Bike	*Skater*	*Cart*	*Walk Across*	*Other*	Total[b]
Ped1	34	19/28	13/13	6/6	3/4	3/3	36/54

Table 1: UCSD Dataset

Figure 2: Biker

Multiple object tracking using OpenCV decision tree classifier in machine learning involves several steps. First, the input frames are preprocessed to extract relevant features such as color, size, and shape. These features are then used to train a decision tree classifier to recognize different objects. Next, object detection is performed on the frames using the trained classifier. Finally, object tracking is carried out using techniques such as Kalman filtering and Hungarian algorithm. The performance of the tracking algorithm is evaluated using metrics such as tracking accuracy, precision, and recall. Overall, the OpenCV decision tree classifier is a powerful tool for multiple object tracking, as it can handle complex datasets and produce accurate results. We got results frame wise of precision around 0.80 and recall of 0.15 .

Figure 3: Truck

Figure 4: Car

V. CONCLUSION AND FUTURE WORK

In conclusion, this paper presents Multiple objects tracking framework for video sequences using optical flow features and machine learning techniques. We proposed extracting histogram of oriented optical flow features, training a Decision Tree classifier in an unsupervised manner and detecting anomalies at different detection intervals. The proposed approach is evaluated on the benchmark UCSD Ped1 dataset, achieving promising results in terms of precision, recall, F1 score. While the current framework demonstrates the efficacy of unsupervised anomaly detection in videos, there are opportunities to improve the performance further using deep learning models, 3D spatio-temporal features, advanced clustering algorithms and semi-supervised learning. Additionally, to improve MOT using machine learning, some possible directions for future work include Incorporating attention mechanisms to focus on relevant regions in the image and reduce computational overhead, developing new deep learning architectures that can handle occlusions and track objects more accurately, Developing real-time.

VI. REFERENCES

[1] V. Mahadevan, W. Li, V. Bhalodia and N. Vasconcelos, "Anomaly detection in crowded scenes," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 2010, pp. 1975-1981, doi: 10.1109/CVPR.2010.5539872.

[2] Xin Li Kejun Wang,Wei Wang and Yang Li Engineering Training Center of HarBin Engineering University Automation College of Harbin Engneering University Harbin, Heilongjiang Province, 150001, China "A Multiple Object Tracking Method Using Kalman Filter" 2013

[3] M. Jaward, L. Mihaylova, N. Canagarajah and D. Bull Department of Electrical and Electronic Engineering, University of Bristol Woodland Road, Bristol, BS8 IUB, UK " Multiple Object Tracking Using Particle Filters" 2015

[4] Video anomaly detection and localisation based on the sparsity and reconstruction error of auto-encoder M. Sabokrou, M. Fathy, M. Hoseini

[5] A Review Ratheesh Ravindran, Member, IEEE, Michael J. Santora,Member for "Multi-Object Detection and Tracking, Based on DNN, for Autonomous Vehicles: IEEE, 2021.

[6] Christian Szegedy Alexander Toshev Dumitru Erhan "Deep Neural Networks for Object Detection "2013.

[7] Duraipandian, Kavitha. (2021). Multiple Object Recognition Using OpenCV. Revista Gestão Inovação e Tecnologias. 11. 1736-1747. 10.47059/revistageintec.v11i2.1795.

[8] Manisha R. Nangre Electronics and telecommunication Department Sinhgad institute of technology and science, Pune University Pune-41 " A Survey: On Multiple Object Detection and Tracking "India International Journal of Engineering Research & Technology (IJERT) 2019

[9] A. Brdjanin, N. Dardagan, D. Dzigal, and A. Akagic, "Single object trackers in opencv: A benchmark," in 2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA). IEEE, 2020.