IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Object Tracking Using Python

1Kanduri Ashish, 2Bairy Yeshwanth, 3Lagadu Gopi Chand, 4Naga Mallikarjuna
1Student, 2Student, 3Student, 4Student
1Lovely Professional University,
2Lovely Professional University,
3Lovely Professional University,
4Lovely Professional University

ABSTRACT: OpenCV object tracking is a widely used technique in the field. A variety of object tracking-specific features are already incorporated into OpenCV. MediandFlow and MIL are some of the object trackers in OpenCV. Tracks the path taken by an object in a movie using an Object Tracking System. We're detecting objects in movies and webcam images with Python and the OPENCV module in this project. "Browse system videos" and "Start webcam video tracking" are two modules included in this application. Working with methods such as frame differencing, color-space transformation, background separation, optical flow, and a classifier based on the Haar cascade, the project entails implementing numerous object recognition and tracking techniques in video. These approaches include: 1. Besides these methods, a widely used and highly effective edge detection method is also used. Python is used for all of the implementations. The results are extensive, and they are thoroughly evaluated.

Opencv and python are two of the most often used terms.

1. INTRODUCTION

Tracking an object when there is a lot of variation is extremely difficult. Background movement, partial and complete occlusions of complex-shaped objects, and varying degrees of illumination One of the most crucial applications for industries to ease the user, save time and achieve parallelism is object detection and localization in digital images. A more efficient and precise method of object detection is still needed in order to attain the desired result. The primary goal of computer vision research and study is to design a system that reduces human effort by showing the fundamental block diagram of detection and tracking by using a computer. Detection and tracking are implemented in a python environment using SSD and Mobile Nets-based techniques. Object detection is the process of identifying an object's specific area of interest in a certain type of image. Frame differencing, optical flow, and background subtraction are a few of the ways that can be used. With the use of a camera, this is a way to track down and locate an object in motion. By extracting the properties of images and videos for security applications, detection and tracking methods are explained.

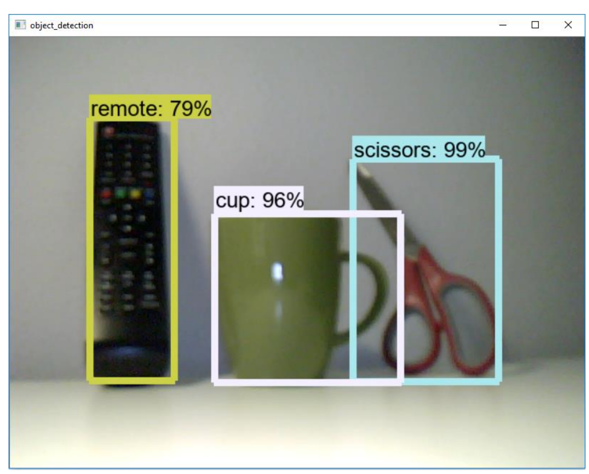

Figure:1. Example figure

2. LITERATURE SURVEY

2.1 Single Shot MultiBox Detector:

A single deep neural network is used to recognise objects in photos. The output space of bounding boxes is discretized into a series of default boxes with varied aspect ratios and scaling depending on the feature map location. We call this approach SSD. At prediction time, the network generates scores for the existence of each item category in each default box and creates updates to the box in order to better match the object's shape. To accommodate objects of varying sizes, the network uses numerous feature maps with varied resolutions to forecast their location. Methods that require a proposal generation stage and subsequent resampling of pixels or features are more complicated than our SSD model, which does not require any proposal generation or resampling. A detecting component is not required, therefore SSD is a snap to learn and integrate into systems. SSD's performance on the PASCAL VOC, MS COCO, and ILSVRC datasets is comparable to those of approaches that use an additional object proposal phase and is substantially faster, while providing a unified framework for both training and inference. SSD's accuracy is superior to that of other single-stage approaches, even when the input image size is less. With 300x300 input, SSD is capable of attaining 72.1% MAP on the VOC2007 test, while with 500x500 input, SSD is capable of attaining 75.1% MAP, exceeding the Faster R-CNN model.

Convolutional Neural Networks for Mobile Vision: 2.2 MobileNets.

For mobile and embedded vision applications, we offer a new class of models termed MobileNets. In order to construct lightweight deep neural networks, MobileNets rely on a simplified architecture that makes extensive use of depth-wise separable convolutions. It is our goal to introduce two basic global hyper-parameters that efficiently balance off latency versus precision. Using these hyper-parameters, the model builder can pick the appropriate model size based on the constraints of the task. Extensive studies on resource and accuracy trade-offs are presented, and we demonstrate high performance on ImageNet classification when compared to other popular models. A wide range of applications and use cases, such as item detection, finegrain categorization, face attributes, and large-scale geo-location, are then demonstrated to show the usefulness of MobileNets.

Gun Detection Classifier Development:

Real-time object detection using Convolutional Neural Networks is a promising application (CNNs). Detection of hand-held weapons is one example of a use case (such as pistols and rifles). Until now, most of the research has concentrated on detecting concealed weapons using infrared data. When it comes to weapons detection and identification, we're more concerned in the speed of the process. A Tensorflow-based implementation of the Overfeat network was used in our study to recognize and classify firearms in photos. 93 percent training accuracy and 89 percent test accuracy were attained using hyperparameter adjustments on Overfeat-3.

Unmanned aerial vehicles (UAVs) can use object detection and tracking techniques.

Unmanned Aerial Vehicle (UAV) data is used extensively in this study to track and locate moving items that pose a significant security concern to the US southern border. For the United States Border Patrol and DHS, dealing with border trespassers and people trying to cross illegally is a major challenge. Due to the enormous volume of data involved, it becomes impossible to guarantee suspicious behavior through long-term monitoring by human operators. Ultimately, this study hopes to aid human operators by developing intelligent visual surveillance systems that can recognize and track unusual or suspicious events in video sequences. There must be efficient ways of recognizing and tracking moving objects in the visual surveillance system. In this study, we looked into ways to detect and track things from unmanned aerial vehicles. Adaptive background subtraction was used to successfully detect moving objects, and these detected items were tracked using Lucas-Kanade optical flow tracking and Continuously Adaptive Mean-Shift tracking techniques. UAV video sequences captured by the UAV reveal that these methods are effective at detecting and tracking moving objects.

Deep learning and OpenCV are used to detect objects in images.

Developing techniques for recognizing images and displays is the goal of computer vision. In addition to image recognition and object detection, it can also create new images. Face detection, vehicle detection, web photos, and safety systems all make use of object detection in some way. As a goal, the YOLO (You Only Look Once) method is used to discriminate between items. When compared to other object detection methods, this method has a few distinguishing features. When compared to other algorithms, such as Convolutional Neural Network and Fast-Convolutional Neural Network, YOLO looks at the image more thoroughly because it uses convolutional networks to anticipate the boundaries of boxes and the class probabilities associated with those boxes, allowing it to identify the image more quickly. Deep learning algorithms, which are also based on machine learning, necessitate the use of several mathematics and deep learning frameworks, such as OpenCV, in order to detect and identify each individual object in a picture by the area object within a highlighted rectangular box. This also takes into account the precision of each method of object differentiation.

2.6: Video Surveillance Object Tracking Algorithms:

Object tracking is one of the most important fields of research because of the constant movement of the object. Object tracking relies heavily on feature selection. For example, it can be used for a wide range of tasks including object detection and traffic control, as well as human-computer interaction (such as gesture recognition and video monitoring). Smoothing the video sequence is a common way to dealing with tracking concerns such as object movement and appearance. For video surveillance applications, these tracking systems use object shape, color, texture, item of interest, and motion in several directions. An comprehensive investigation of numerous object tracking algorithms under a variety of environmental conditions was carried out in this research, which discovered an efficient approach for a variety of tracking applications. Single and multiple vehicle motion is recognized and counted in numerous frames in this work based on color-based tracking. In addition, a single tracking algorithm can be built that takes into account the object of interest's shape and color as well as its mobility in many directions.

Using Convolutional Neural Networks (CNNs) and Deep Learning to extract features:

One of the first things we learn as children is how to classify images. Images can be classified by determining the basic shapes and geometric properties of the items in our environment. Pre-processing the image (normalization), segmentation, feature extraction, and class identification are all part of the process. With today's faster and more accurate picture classification techniques, they may now be utilized for a wide range of applications, including security features, face recognition for authentication and authorisation, traffic identification, medical diagnostics, and many more. The concept of picture classification can be approached in a variety of ways. Machine learning algorithms are the most effective. A similar concept was offered years ago but was never put into practice due to a lack of computing power. Models are better trained and able to recognize multiple levels of visual representation thanks to the concept of deep learning. The convolutional neural networks changed this discipline by learning the basic shapes in the initial layers and growing to learn aspects of the image in the deeper layers, which resulted in more accurate classification of images. For the inspiration for the Convolutional neural network, Hubel and Wiesel in 1962 developed a hierarchical representation for neurons that was based on the research of the visual cortex in a cat. In the realm of computer vision, it was a major breakthrough in understanding how the visual cortex works in humans and animals. In this paper, a convolution neural network is used to extract a feature from a picture using deep learning. For a variety of applications, additional categorization methods are implemented.

A method that uses YOLO to detect and classify objects in video recordings (2.8)

Each problem is broken down into smaller modules and solved individually by the primitive machine learning techniques currently in use. A detection algorithm must be able to work from beginning to end and in the shortest amount of time possible in order to be effective nowadays. Many analytical features can be generated from real-time identification and categorization of objects from video records, such as the volume of traffic in a certain area over time or the overall population in an area. Due to the limited and lightweight datasets used, classification and detection can take a long time or lead to incorrect detections in practice. With these difficulties in mind, we've developed an improved version of the 'You Only Look Once' detection and classification approach (YOLOv2), which is able to increase computing and processing performance while also efficiently classifying objects in video records. Every class of object for which the classification algorithm has been taught has a bounding box and an annotation indicating the specific type of object.. A GPU (Graphics Processing Unit) is used in YOLOv2's detection and classification, allowing it to process at a rate of 40 frames per second.

3. PROPOSED SYSTEM

We're detecting objects in movies and webcam images with Python and the OPENCV module in this project. "Browse system videos" and "Start webcam video tracking" are two modules included in this application.

Allows the user to upload a video from his system and the application will begin to play that video; if the application detects any objects, those objects will be marked with bounding boxes. If the user wishes to stop tracking the video while it is playing, he or she can press the "q" key on their keyboard to do so.

In order to use this module, the program must first connect to the built-in system webcam and begin streaming video; if the application detects an object during this process, the application will surround that object with bounding boxes. To stop the webcam streaming, simply hit 'q.'

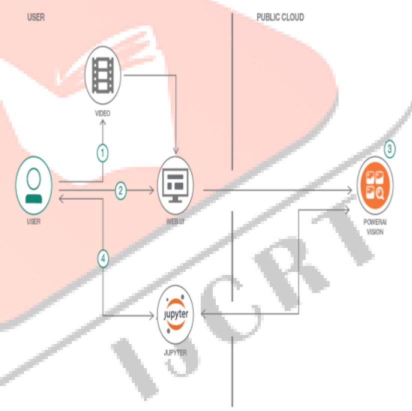

Figure.3: System architecture

OBJECTIVES:

Object tracking systems are designed to improve the detection and tracking of objects. contributing to the development of two distinct components in the beginning

2) Object tracking and motion segmentation

Here are our top priorities:

To locate a moving object's coordinates

For feature extraction and classification, a GMM model with a YOLO-based method was analyzed.

OpenCV may be used to determine how fast an object is moving in a video.

To study the tracking algorithms of SSD and Mobile Nets

The OPENCV python API's object tracking techniques were used to create this project.

What does the term "object tracking" refer to?

"Tracking" is a term that refers to the process of locating a moving target in successive frames of video. tracking is a phrase that covers a wide range of concepts that are both conceptually related and technically distinct in computer vision and machine learning. Object tracking, for example, encompasses a wide range of topics that are all connected but study separately.

Dense Optical Flow: Opencv will employ the following methods to track moving objects in videos:

All pixels in a video frame are given an estimated motion vector thanks to these techniques. There is a technique known as "sparse optical flow" that uses a small number of feature points in an image to figure out where things are in the picture.

Use of a Kalman Filter:

Predicting the location of a moving item using prior motion information is a common signal processing approach. This algorithm was used in the early stages of missile guiding! A Kalman filter was used by the Apollo 11 lunar module's onboard computer, which is also referenced in this article.

Camshift and Meanshift

A density function's maxima can be found using one of these algorithms. Their usage in tracking is another benefit of using them.

Trackers for single objects:

Using a rectangle in the first frame of this type of tracker, we may identify the object we wish to follow. The tracking method is then used to keep track of the item in successive frames. These trackers are typically used in conjunction with an object detector in most real-world applications.

INTERACTIVE TRACKING AND DETECTION

Face detection in OpenCV works in real time, allowing you to identify the face in any given frame. Since tracking is so important, what is its purpose? Instead of simply performing repeated detections, consider tracking moving objects in a movie for a moment.

Detection is slower than Tracking:

Detection methods are typically slower than tracking systems. It's easy to see why. It's easier to track an object that was previously spotted since you have a good idea of how it will appear. You also know where the object was in the previous frame, as well as its speed and direction of travel. So in the following frame, you can use this information to estimate the location of the object and conduct a brief search around the expected location of the object in order to locate the object more correctly. While a detection algorithm always starts from scratch, a smart tracking program uses all the knowledge it has about the object up to that moment. As a result, when developing an effective system, it is common practice to run an object detection algorithm every nth frame, followed by a tracking method every n-1 frames. Detect the object in the first frame, and then follow it as it moves around the screen. You can, however, lose track of an object if it is buried behind an impediment for a long length of time or if it moves so quickly that the tracking algorithm is unable to keep up. An object's bounding box might also drift away from it over time due to tracking algorithms accumulating errors. Every now and again, a detection algorithm is run to remedy the issues with tracking algorithms. A vast number of examples of the object are used to train detection algorithms. To put it another way, they have a better grasp on the object's general classification. tracking algorithms on the other hand are more knowledgeable about the specific instance of the class they are tracking. When detection fails, tracking can come in handy. If an object obscures a person's face while running a face detector on a video, the face detector is likely to fail. In contrast, a competent tracking algorithm can handle some occlusion.

4. EXPERIMENTAL RESULTS

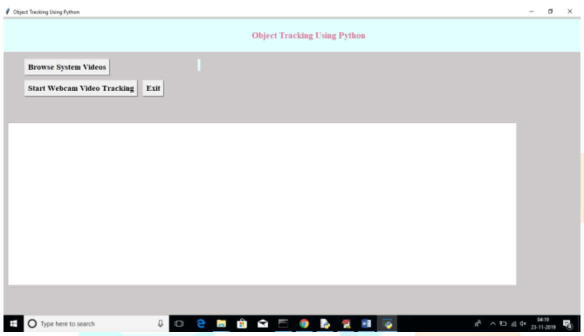

Fig.4: Home screen

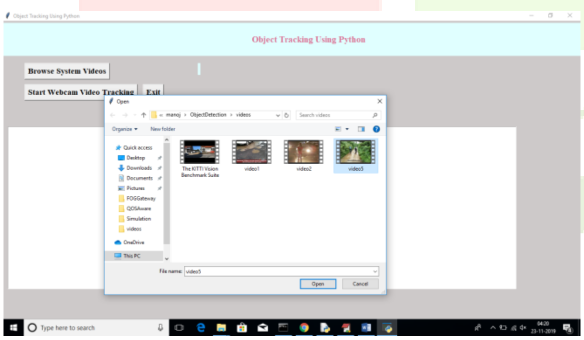

Fig.5: Videos screen

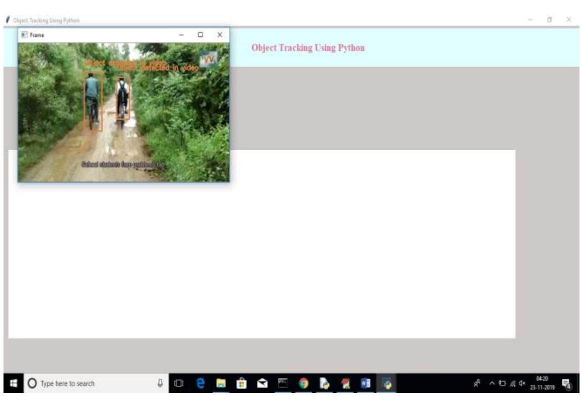

Fig.6: Start tracking objects from video and mark them with bounding boxes

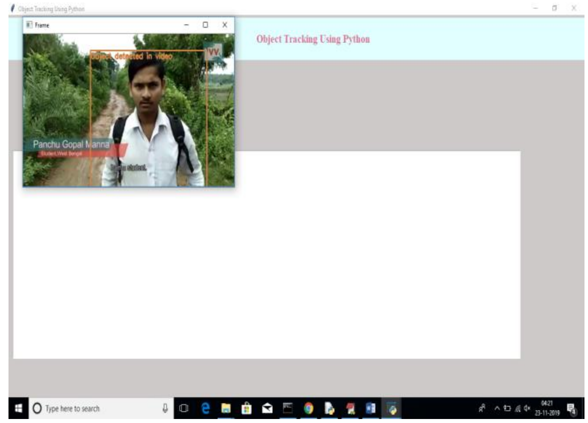

Fig.7: Result from video

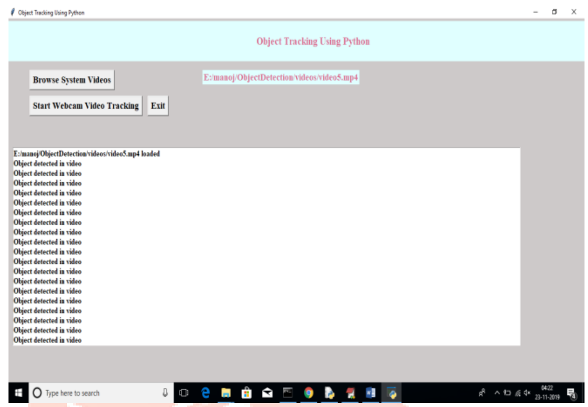

Fig.8: Object detected

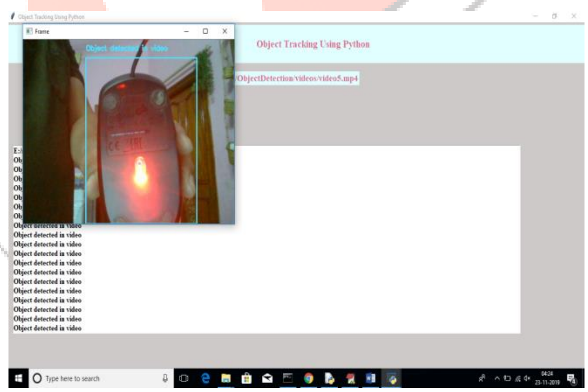

Fig.9: Start webcam video tracking

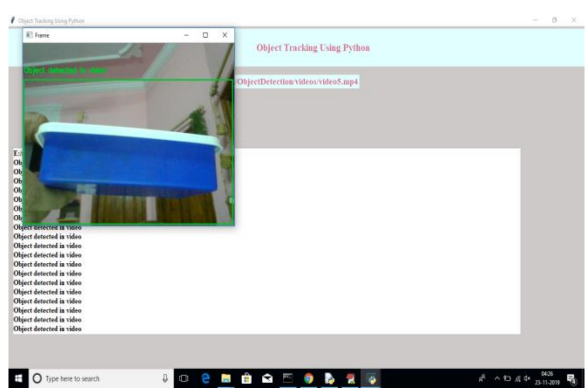

Fig.10: Object detected in video

5. CONCLUSION

The current state-of-the-art object detection technology has been replaced by a new system that is both accurate and efficient. OpenCV is used for object detection in this project, which employs Python and the OPENCV package.

6. FUTURE ENHANCEMET

It's possible to make the visual tracking method provided here more robust by removing some of the following limitations: When using Single Visual tracking, the template's size remains constant. When the size of the monitored object shrinks with time, the background takes on more importance. The object may not be able to be tracked in this situation. There can be no tracking or new object consideration for an object that is completely obscured. Multi-view tracking with many cameras may be used in the future to automate the system and overcome the constraints listed above. In comparison to single view tracking, multi view tracking has an evident advantage due to its ability to cover a much larger area with a wide range of viewing angles for the tracked objects.

REFERENCES

1. Wei Liu and Alexander C. Berg, SSD: Single Shot MultiBox Detector, Google Inc., Dec 2016.
2. Andrew G. Howard, and Hartwig Adam, MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, Google Inc., 17 Apr 2017.
3. Justin Lai, Sydney Maples, Ammunition Detection: Developing a Real-Time Gun Detection Classifier, Stanford University, Feb 2017
4. ShreyamshKamate, UAV: Application of Object Detection and Tracking Techniques for Unmanned Aerial Vehicles, Texas A&M University, 2015.
5. Adrian Rosebrock, Object detection with deep learning and OpenCV, pyimagesearch.
6. Mohana and H. V. R. Aradhya, "Elegant and efficient algorithms for real time object detection, counting and classification for video surveillance applications from single fixed camera," 2016 International Conference on Circuits, Controls, Communications and Computing (I4C), Bangalore, 2016, pp. 1-7.
7. AkshayMangawati, Mohana, Mohammed Leesan, H. V. Ravish Aradhya, Object Tracking Algorithms for video surveillance applications International conference on communication and signal processing (ICCSP), India, 2018, pp. 0676-0680.
8. ApoorvaRaghunandan, Mohana, PakalaRaghav and H. V. Ravish Aradhya, Object Detection Algorithms for video surveillance applications International conference on communication and signal processing (ICCSP), India, 2018, pp. 0570-0575.
9. ManjunathJogin, Mohana, Feature extraction using Convolution Neural Networks (CNN) and Deep Learning 2018 IEEE International Conference On Recent Trends In Electronics Information Communication Technology,(RTEICT) 2018, India.
10. ArkaPrava Jana, AbhirajBiswas, Mohana, YOLO based Detection and Classification of Objects in video records 2018 IEEE International Conference On Recent Trends In Electronics Information Communication Technology,