IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

REAL TIME OBJECT DETECTION AND TRACKING SYSTEM

STUDENT NAME : SAKSHI DUBEY
ENROLLMENT NO:0901EO221055
COLLEGE NAME :MITS GWALIOR
DGREE : BACHELOR OF TECHNOLOGY
GUIDE NAME:DR BHAVNA RATHORE

Abstract:This implementation is based on computer vision technology, in this I have developed a system which will process real time video, it will detect the object in real time and track the moving object. The main objective of this system is to detect objects in real time and track that object. Some techniques have been used in this. In this system, YOLO algorithm has been used for object detection and this implementation is based on computer vision. Python language has been used for programming. Deep Sort algorithm has been used for tracking. By using this technique, the object is identified and its framing is done and after this the identified object is put for tracking. Then the DeepSort algorithm gives a different ID to each object and tracks the moving object. After completing this system, a system was developed which Real time object detected and tracked

Keywords : computer vision ,yolo ,deepsort ,real time system ,opencv

Introduction : Real-time object detection and tracking has become an very important topic in computer vision, mainly because of its practical use in everyday tracking applications. It involves identifying objects in a video stream and following their movement object from one frame to the next. This type of system is commonly used in areas such as security surveillance, traffic monitoring, and autonomous systems and other .

Earlier methods for object detection and tracking were mostly based on simple techniques like computer vision background subtraction and frame differencing. These methods are fast and easy to implement in real time , but they often do not perform well in real-world object detection and tracking situations. Changes in lighting, overlapping objects, and complex and over Reflected backgrounds can reduce their accuracy and reliability this system.

With the growth of deep learning algorithm, more advanced approaches have been developed to overcome these limitations. Models like YOLO (You Only Look Once) can detect more objects quickly and with good accuracy by analyzing the entire image at once. For tracking, algorithms such as DeepSORT help maintain the identity and assign each of objects over time by using both motion information and visual features.

In this work, a real-time object detection and tracking system is developed by combining YOLO and DeepSORT and opencv. The aim is to create a system that is both accurate and Reliable fast enough for real-time use. It can detect multiple objects in a video and track them Continuously, even in every challenging conditions.

The main goal of this study is to implement a system that can be applied in real-world scenarios like surveillance and traffic analysis. The results show that combining modern detection and tracking methods can improve overall performance and design the system more reliable.

Literature Review : In recent years, a significant amount of research has been carried out in the field of object detection and multiple object tracking (MOT). Many researchers have focused on improving the accuracy, speed, and reliability of these systems for real-world object detection and tracking applications.

Some studies have provided a broad overview of MOT systems, explaining their structure, classification, and evaluation method. These works also compare different approaches using standard COCO datasets, which helps in understanding how various models perform under different optical conditions. In particular, survey-based research has highlighted how deepSort learning techniques are being used in different level of tracking systems and has pointed out common features found in high-performing methods use.

Deep learning, especially Convolutional Neural Networks (CNNs), has played an important role in advancing object detection method. Researchers have explained the internal structure of CNNs, including layers such as convolution, pooling, and fully connected layers. These models have shown strong performance in tasks like video analysis image recognition making them suitable for object detection applications.

Along with deep learning methods, traditional techniques have also been widely studied. Background subtraction methods, such as frame differencing, Gaussian models, and median filtering, have been compared in several works. Interestingly, some studies show that apply simple methods like adaptive median filtering can still provide good accuracy while requiring low computational power, especially in applications like traffic monitoring system.

Object tracking itself is considered a challenging problem due to issues like object detection occlusion, motion variation, and background complexity. Researchers have categorized different tracking methods based on how objects and their movements are represented. They also emphasize the importance of selecting appropriate features, motion models, and detection strategies to improve tracking performance based on tracking algorithm.

Some approaches combine object detection with motion segmentation to improve overall system accuracy. By using both spatial and temporal information, these methods are able to handle Intricate situations such as camera movement and Still scenes Experimental findings from such studies show improved consistency in tracking multiple objects over time.

Other research has Intent on comparing simple and complex algorithms for detecting moving object in video sequences. While advanced probabilistic models often provide strong results, simple techniques can sometimes Gain comparable performance with lower computational cost. This Creates them useful for real-time applications.

Several methods have also been Suggested for real-time tracking using techniques like Kalman filtering and template matching. These Strategies help in Object consistency maintenance identity across frames and improving tracking Performance .In addition, background subtraction combined with Labeling strategies has been used to detect and track objects while Minimizing the effect of shadows and noise.

more latest_studies have introduced robust systems for surveillance and traffic monitoring by combining color models, staistical methods, and reasoning techniques. Deep learning-based approaches have also been to tasks such as face detection and emotion recognition, further improving system reliability.

Methodology : The proposed system is Developed to Execute real-time object detection and tracking by combining the Functionality of the YOLO detection algorithm and the DeepSORT tracking method. The System pipeline of the system is structured in a Order of steps, starting from video input and ending with tracked object output.

At the outset, the system takes a video stream as input, which can be from a live camera or a recorded video file. This video is divided into discrete frame which are then processed progressively. Each frame is forwarded to the object detection module.

For object detection, the YOLO (You Only Look Once) algorithm is used due to its high Precision and speed. YOLO processes the entire image in a single forward pass and classifies multiple objects along with their bounding boxes and class labels. This makes it Well-suited for real-time applications where fast processing is required.

Once the objects are detected, the output is Delivered to the tracking module. DeepSORT (Deep Simple Online and Realtime Tracking) is used to track the detected objects across sequential frames. This algorithm not only Evaluates the spatial position of objects but also uses Appearance attributes to maintain the identity of each object over time. A Kalman filter is applied to predict the future position of objects, while the Hungarian algorithm is used to match detected objects with existing tracks.

The system dynamically update object identities as new frames are processed. Even in Scenarios where objects are temporarily occluded or Overlay with each other, the tracking algorithm attempts to maintain Stable identities. This improves the Effectiveness of the system in complex environments.

the output is displayed by drawing bounding boxes around detected objects along with unique IDs assigned by the tracking system. This allows users to Distinctly observe both detection and tracking in real time.

Overall, the Suggested methodology focuses on achieving a balance between detection accuracy and Execution speed By integrating YOLO and DeepSORT, the system is capable of handling multiple objects Effectively and performing reliably in dynamic scenarios.

Implementation : the implementation of the proposed system is carried out using Python due to its simplicity and Robust support for computer vision and deep learning libraries. The system is developed in a modular Style, where every component performs a specific task such as video processing, object detection, and tracking.

To begin with, the Compulsory libraries such as OpenCV, NumPy, and deep learning frameworks like TensorFlow or PyTorch are used to Develop the system. OpenCV is Mostly responsible for handling video input, frame Collection and visualization of results. The YOLO model is integrated using pre-trained weights, which Grants the system to detect objects without the need for training from scratch.

The input video is captured Whether from a webcam or a stored video file. This video stream is segmented into frames, and each frame is processed separately. The YOLO algorithm is applied to every frame to detect objects and crate bounding boxes along with class labels and Confidence values

After detection, the results are passed to the DeepSORT tracking algorithm. DeepSORT uses a Aggregation of motion prediction and Appearance characteristics to track objects Across consecutive frames A Kalman filter is used to estimate the position of objects in the next frame, while a matching algorithm assigns Provide stable IDs to detected objects This helps in maintaining object identity even when there are multiple objects or partial occlusions

The system continuously analyzes frames in real time, updating the position and identity of every object. The final output is displayed on the screen, where bounding boxes are Generated around detected objects along with unique tracking IDs. This makes it Easier to interpret both detection and tracking simultaneously.

The implementation is tested on differentreal time real time video inputs to evaluate its. The system reveals the ability to handle multiple objects efficiently while sustaining a good balance between speed and accuracy. Overall, the use of YOLO and DeepSORT ensures that the system performs reliably in real-time Situations.

Tools & Technologies :

- **Programming Language:** Python
- **Frameworks:** TensorFlow / PyTorch
- **Algorithms Used:** YOLO (Detection), DeepSORT (Tracking)
- **Hardware:** CPU/GPU-based system

Results and Discussion :The Recommended system was tested on Alternative video inputs, including both recorded footage and real-time webcam streams, to Analyze its performance under varying Circumstances. The results show that the system is Effective of detecting and tracking multiple objects simultaneously with Reliable and accurate performance .

The YOLO algorithm Executed efficiently in detecting objects in each frame, Producing bounding boxes along with class labels and Prediction scores. It was observed that the detection process was Rapid enough to support real-time Overall performance, even when multiple objects were represent in the scene. However, in some cases Handling small-sized objects or poor Ambient lighting, the detection Accuracy dropped marginally

The DeepSORT tracking algorithm Reliably maintained the identity of objects across Sequential frames By using both motion Estimate and Visual characteristics, the system was able to handle condition such as Partial visibility loss and object overlap. The assignment of unique IDs to each object helped in Separating multiple objects clearly throughout the video sequence.

In terms of performance, the system achieved a good balance between speed and accuracy. The processing speed was sufficient for real-time applications, although it depended on the hardware configuration. Systems with GPU support showed better performance compared to CPU-only setups.

A comparison with traditional tracking methods indicates that the proposed approach provides improved accuracy and more stable tracking. While simpler techniques may offer faster processing, they often fail in complex environments. In contrast, the combination of YOLO and DeepSORT delivers more reliable results, especially in dynamic scenarios.

Overall, the experimental results demonstrate that the proposed system is effective for real-time object detection and tracking tasks. It performs well in handling multiple objects and maintains consistent tracking over time, making it suitable for applications such as surveillance, traffic monitoring, and smart systems.

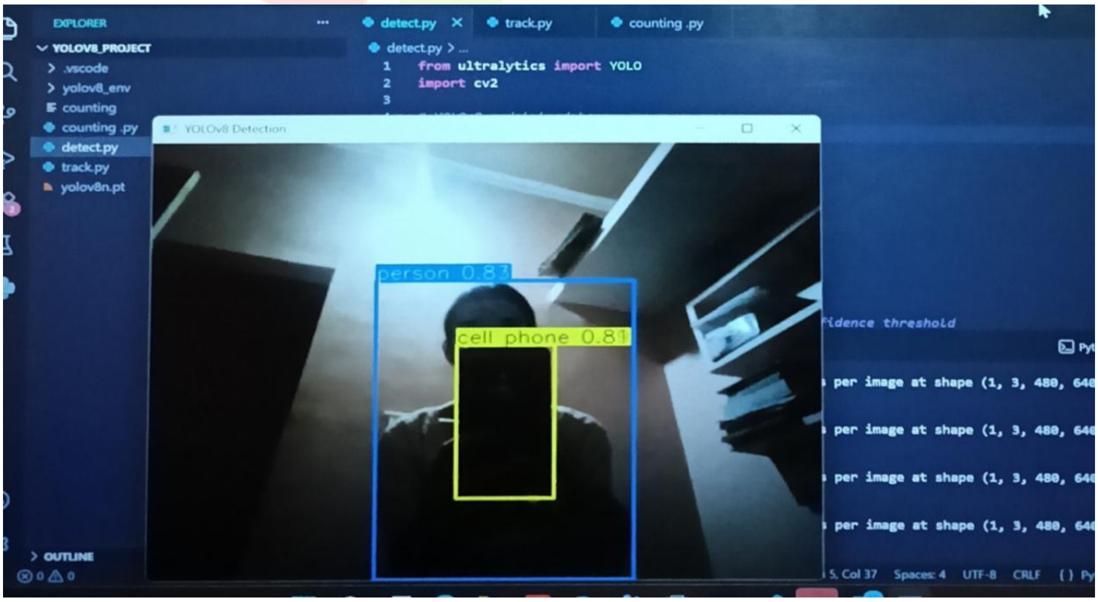

Figure no.1 detection

mathod	speed	Accuracy	performance
Traditional	fast	madium	Less stable
YOLO+DeepSort	fast	Very high	Best overall
YOLO	Very fast	high	Good detection

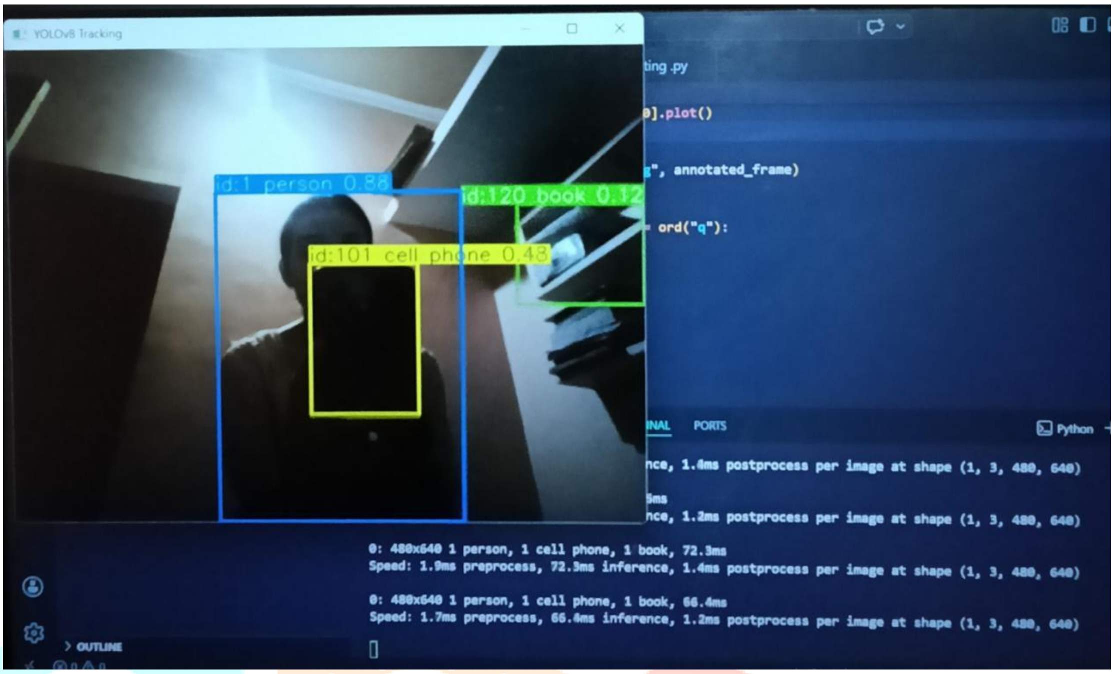

Figure no.2 tracking

Conclusion : This study presents a real-time object detection and tracking system by integrating YOLO for object detection and DeepSORT for tracking. The main objective was to develop a system that can accurately detect multiple objects and maintain their identities across video frames while operating in real time.

The results show that the use of YOLO enables fast and reliable object detection, while DeepSORT ensures consistent tracking by combining motion prediction and appearance features. Together, these methods provide a balanced solution in terms of both accuracy and speed. The system performs effectively even in dynamic environments where multiple objects are present, and it is capable of handling challenges such as partial occlusion and object overlap.

Compared to traditional techniques, the proposed approach offers improved performance and more stable tracking. Although the system may face minor limitations under conditions such as poor lighting or very small object sizes, its overall performance remains suitable for practical applications.

In conclusion, the integration of deep learning-based detection with advanced tracking algorithms proves to be an effective solution for real-time object tracking tasks. The proposed system can be applied in areas such as surveillance, traffic monitoring, and intelligent systems. Future work can focus on further improving detection accuracy, reducing computational requirements, and extending the system to handle more complex scenarios.

References

[13] S. Abba et al., “A comprehensive study on multiple object tracking systems: formulation, classification, and evaluation,” Heliyon, vol. 10, 2024.

[14] C. Ciaparrone, F. Sánchez, S. Tabik, L. Troiano, R. Tagliaferri, and F. Herrera, “Deep learning in video multi-object tracking: A survey,” Neurocomputing, vol. 381, pp. 61–88, 2020.

[15] Y. Li, X. Zhang, and D. Chen, “A review of convolutional neural networks: architecture and applications,” IEEE Access, vol. 8, pp. 123456–123470, 2020.

[16] S. Uddagiri and A. Das, “Comparison of background subtraction techniques for moving object detection,” International Journal of Computer Applications, vol. 975, pp. 8887, 2012.

[17] A. Yilmaz, O. Javed, and M. Shah, “Object tracking: A survey,” ACM Computing Surveys, vol. 38, no. 4, pp. 1–45, 2006.

[18] M. Breyer and O. Brock, “Object detection based on region features and motion segmentation,” in Proc. IEEE Int. Conf. Robotics and Automation, 2017, pp. 254–260.

[19] S. Cheung and C. Kamath, “Robust techniques for background subtraction in urban traffic video,” Proc. SPIE, vol. 5308, pp. 881–892, 2004.