Moving Object Detection using Intelligent Spectaculars

Nachiket Chaudhari[1], Vina.M.Lomte[2], Siddhi Kulkarni[3], Dimpal Patil[4], Neeraj Pawar[5]
Department of Computer Engineering, First-Fifth SPPU

Abstract: Moving Object detection systems are used to identify and locate objects in images or videos. When used spectacles, an object detection system can allow the user to see information about the objects in their field of view. This can be useful in a variety of applications, such as helping blind people navigate their environment, or providing augmented reality information to workers. Region-based Convolutional Neural Networks (RCNN) is a type of machine learning model that is commonly used for object detection. The RCNN model first uses a convolutional neural network (CNN) to extract features from the input image, and then applies a region proposal algorithm to identify potential object regions in the image. These regions are then fed into a second CNN, which classifies the regions as objects or background. The RCNN model has been shown to be effective at detecting a wide range of objects in images and videos.
Keywords: Moving Object Detection Systems, Machine Learning, Region-based CNN, Spectacles, Algorithm.

I. INTRODUCTION

In today's technologically advanced generation, artificial intelligence is something that is constantly changing. Artificial intelligence (AI) is constantly being incorporated into new technologies that improve quality of life, safety and security, entertainment, and many other areas. Artificial intelligence may be permitted to make decisions and learn on its own because it was built with the explicit goal of becoming automated and automating anything to which it is applied. Object identification is one such application of AI that helps with safety and quality of life.

The emergence of new media over the past few years has drastically altered people's behaviour, with smart glasses serving as a notable example. The goal Intelligent Spectacular for Blind People has always been to reduce the distance between the physical and digital worlds.

The manner in which humans interact with mobile devices via wireless communication. Individuals with visual impairments and blindness have difficulty moving independently and overcoming various problems in their daily lives. As a solution, artificial intelligence and computer vision approaches help blind and visually impaired (BVI) people carry out their primary activities without relying on others.

WIFI and Bluetooth will be used for system communication between devices. The camera module on the wearable Intelligent Spectacular for Blind People will capture Realtime continuous moving images of objects and send them over the device for classification and identification. The captured image will be sent to the cloud server. The cloud server will enhance the image according to its needs and perform object detection and identification, after which the image labelling of the object will be completed and sent back to the transmitting point. "Object recognition is a computer vision technique used to recognise objects in images or videos. Deep learning and machine learning algorithms provide significant amounts of object recognition. The goal of object detection is to aid machine learning by teaching what comes naturally to humans: the capacity to effortlessly identify items in any video or picture, such as people, automobiles, animals, objects, and more. As a result of teaching these things, AI may learn and obtain a level of comprehension of what a picture includes, allowing it to analyse and extract elements from an image.

II. RELATED WORK

Object detection using convolutional neural network: SSD with mobilenetv1 offers high detection speed but low accuracy when compared to faster-run with inceptionv2, which has low detection speed but higher accuracy. According to the findings of the trials, there is a trade-off between accuracy and speed. Use SSD with mobilenetv1 if we desire quick detecting capabilities, especially in real-time applications. Faster-rcnn with inceptionv2 is recommended for high-accuracy detection. In the future, the two models will be used as the vision system of a bomb disposal robot to detect improvised explosive devices (ieds).

Fig. 1 shows an example of a low-resolution image and high resolution image.

Appearance and motion based deep learning architecture for moving object detection in moving camera: Deep learning-based moving object recognition framework for freely moving camera recordings, such as dashcam footage. The suggested technique comprises of two networks: one focused on appearance and the other on mobility. The suggested deep learning technique delivers a strong performance against background contamination, even with unconstrained camera movement, which is a significant advance. The efficiency of the suggested approach was validated in the experiments by comparison with state-of-the-art methods. Furthermore, the suggested technique has a significant benefit in that it can operate at a real-time pace of 50 frames per second, which is suited for real-world applications such as demanding scenarios in autonomous cars.

Moving Object Detection for Event-based Vision using Graph Spectral Clustering: GSCEventMOD is a tool for detecting moving objects based on events using graph spectral clustering. It demonstrated that detecting moving objects with neuromorphic vision sensors can perform well in difficult situations such as fast motions and abrupt changes in lighting conditions. Because scene dynamics are captured by the sensor, GSCEventMOD requires minimal pre-processing. In event-based vision, GSCEventMOD has been shown to outperform some previous approaches. GSCEventMOD was validated using synthetic data and real-world data captured in a variety of environments to demonstrate its adaptability. Overall, it is expected that the method will be applicable to a wide range of computer vision applications, including autonomous vehicles, robotics, remote surveillance, and others.

Moving Object Detection by a Mounted Moving Camera: Finding and tracking interest points in consecutive video frames using the pyramidal Lucas-Kanade method. Then it calculates camera motion based on the assumption that the most common motion vectors belong to camera motion. After the camera motion is removed, the frame difference method is used to detect moving objects. Using adaptive thresholding while taking into account different lighting conditions in the same video.
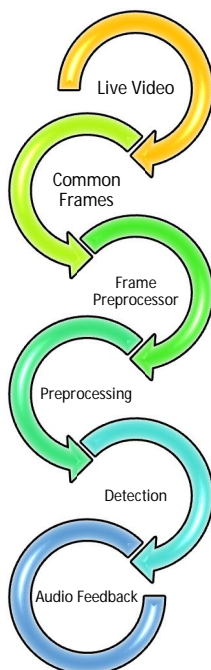
Moving Object Detection and Segmentation using Frame differencing and Summing Technique: When compared to conventional object recognition and segmentation techniques, this approach is straightforward and has a low computational complexity. This technique quickly identifies and segments moving items from the static backdrop. However, because this technique does not consider the shadows of moving objects, they are also segmented as moving objects if they are larger than the cut-off number.

III.   PROPOSED WORK

In this section, we present the details of Camera motion estimation and moving object detection that we are going to use in our system.

There are many blind people in the world. These blind people cannot feel or understand their surroundings because they cannot see. Our invention directly addresses this problem. Our invention, intelligent spectacular for blind people, helps blind people comprehend and identify things through speech. Intelligent spectacular for blind people (wearable goggles) is our innovation that detects and converts continuous moving objects into speech. Computer Vision OpenCV technology is commonly used to capture images and perform some pre-processing methods to reduce the size of the picture to only take usable regions for detection.

A.   Raspberry Pi 3

The Raspberry Pi Foundation, in collaboration with Broadcom, developed a series of compact single-board computers known as the Raspberry Pi. Early on, the Raspberry Pi initiative oriented towards promoting fundamental computer science education in schools and impoverished nations. The Raspberry Pi Foundation, a UK organisation that strives to educate people in computing and make computing education more accessible, created the Raspberry Pi line of single-board computers. Raspberry Pi receives video frame input from the camera and computes the primary algorithm. The Object detection model is delpyed on the Raspberry Pi to detect objects and provide feedback to the user. After capturing continuous video, the background subtraction method is used to detect and classify moving objects. After successfully detecting an object, the output is communicated to the user in the form of audible feedback.

B.   Camer

In our Intelligent Spectacle for Blind People, the camera is mounted on the bridge of the glasses (wearable goggles). Because the camera has 8.0 megapixels, it can record high-quality video pictures. The is connected to the device and is placed in glasses in such a way that it records a broad area of its user's surroundings. The camera first captures video snippets in the form of a series of pictures of moving continuous objects. RCNN begins by labelling an image. It predicts whether the image contains a single or multiple items and then names the class. Object localization in the image is done after the item has been located using a bounding box.

C.   Moving Object Detection

Detecting moving objects can be difficult, but with RCNN and the Raspberry Pi 3 Model B+, it becomes much faster. RCNN, or Region-based Convolutional Neural Network, is a deep learning technique designed especially for image object detection. It works by segmenting the image into regions and then classifying each region as having or not containing an object. This is accomplished by training the model on a big dataset of labelled images. After training, the RCNN model can be implemented on the Raspberry Pi 3 Model B+. Because it has enough processing power to run the RCNN model in real-time, this powerful single-board computer is ideal for this job.

This means that as soon as the camera captures a frame, it processes it and labels the bounding boxes around the items. These bounding boxes can be used to monitor the objects' movement over time. To obtain results, the system will compare these objects to a dataset.

The dataset is made up of images that have been labelled with the objects that can be found in them. This dataset has been used to train the RCNN model, which means it has learned to identify the objects and their features. When the RCNN model is provided with a new image, it analyses it and uses the bounding boxes to recognize the objects. The features of these objects will then be compared to the features of the objects in the dataset. As a result, the model can accurately identify what objects are present in the image. . The interaction of the objects in the image with the dataset is an important stage in the object process. It enables the system to produce accurate and consistent outcomes. Overall, the RCNN and Raspberry Pi 3 Model B+ combination is a powerful tool for identifying moving objects.

D. Headphones

The results of object detection are transferred to the user via audio feedback via headphones in the system outlined previously. This enables the user to receive real-time feedback on the detected objects without having to glance at the screen.Depending on the requirements of the user, the audio feedback system can be designed to provide a variety of information. Based on the detected objects, this information can be used to assist the user in making choices or taking action. Audio feedback is especially helpful in situations where the user must be constantly conscious of their surroundings, such as security or surveillance applications. It enables the user to stay alert and aware of changes in their surroundings without having to continuously look at a screen. Overall, the combination of object detection and audio feedback has the potential to be a powerful tool for a variety of uses. It detects and analyses objects in real time and provides the user with the information they need to make educated choices.

The detection is communicated to the user via audio headphones. The wearable Intelligent Spectacular for Blind People, Camera module will record real-time continuous images of objects and transmit them to the device for categorization and recognition. The captured picture will be sent to a model developed on a Raspberry Pi 3. The picture will then be enhanced to meet its requirements, as well as object recognition and identity, before being labelled and sent back to the transmitting point. The consumer will receive auditory input after identifying an object in a moving picture.

E. Button

The button is linked to the Raspberry Pi via header pins. In the unlikely scenario of a fatal circumstances, the button has been integrated and the blind person can press it. The person whose phone number has been saved receives notification that the blind person is in danger, along with the blind person's present location.

F. Buzzer

An ultrasonic sensor emits a sound pulse in the ultrasonic range. The ultrasonic sensor assists blind persons in navigating their environment by detecting close objects in the form of obstacles and alerting the blind person with a buzzer sound.
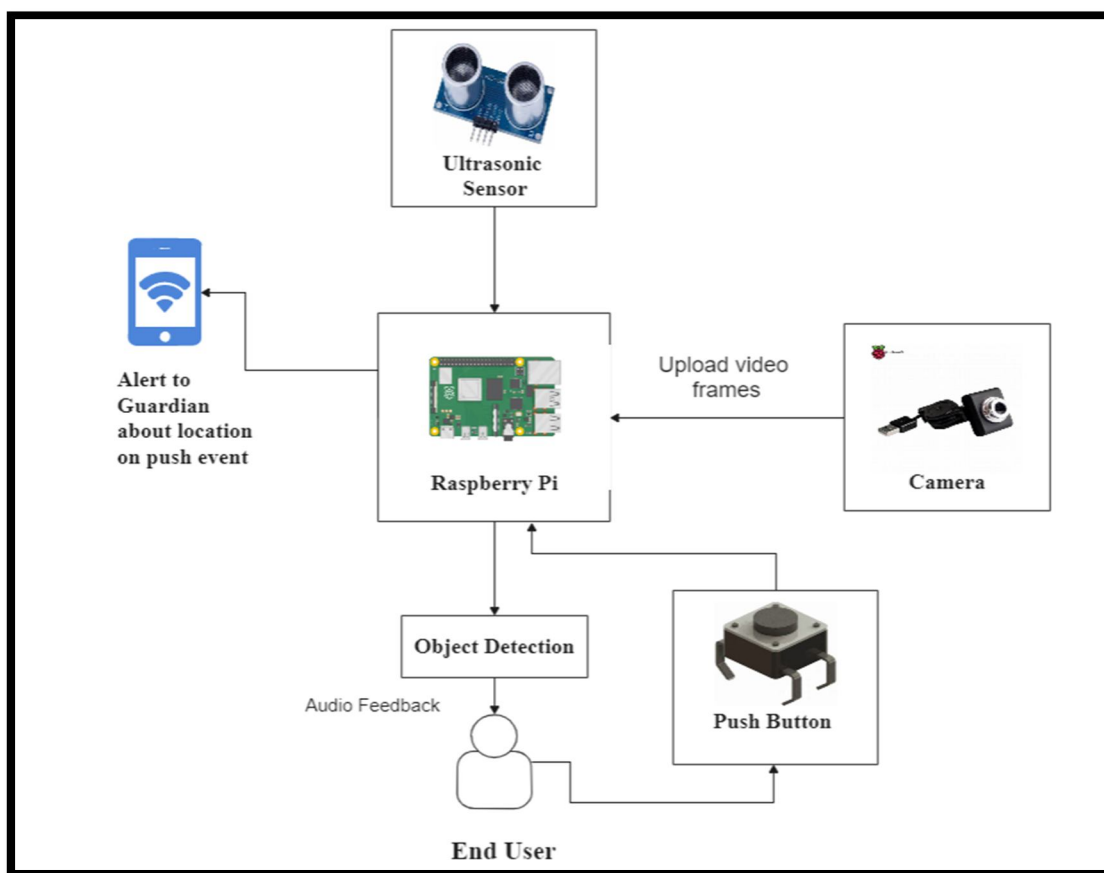

Fig 2. Object Detection Model

IV.   SYSTEM RESULTS

TABLE I

Segmentation(%)	Object Labelling(%)	Moving Object Detection(%)	Audio Feedback(%)
86.3	78	75	75
84	73.2	60	60

V.   CONCLUSION

We present a technique for detecting moving objects in pictures. Like most prior techniques, which are predicated on the premise that each positive picture contains just one object, Image enhancement, motion detection, object tracking, and behaviour understanding studies have all been explored in order to analyse pictures and extract high-level information. In this study, we investigated and presented many strategies for detecting moving objects in cameras. The disadvantage of temporal differencing is that it does not extract all important pixels of a foreground item, particularly if the object has uniform texture or moves slowly. When a moving foreground object stops, the temporal differencing approach fails to identify a change between successive frames and loses track of the object. This paper provides useful insight into this essential research problem and promotes future study in the areas of moving object identification and computer vision. Various estimate approaches are utilised in the kernel tracking strategy to determine the appropriate area to the target item. Nowadays, the most common and recommended kernel tracking approaches are Mean-shift tracking and particle filtering. Contour tracking may be separated into two methods based on how the contours evolve: state space technique and energy function minimization approach.

REFERENCES

[1]   Suresh, Aswath & Arora, Chetan & Laha, Debrup & Gaba, Dhruv & Bhambri, Siddhant. (2019). Intelligent Smart Glass for Visually Impaired Using Deep Learning Machine Vision Techniques and Robot Operating System (ROS). 10.1007/978-3-319-78452-6_10.

[2]   Saha, Himadri & Dey, Ratul & Dey, Shopan. (2017). Low cost ultrasonic smart glasses for blind. 10.1109/IEMCON.2017.8117194.

[3]   E. Shreyas, M. H. Sheth and Mohana, "3D Object Detection and Tracking Methods using Deep Learning for Computer Vision Applications," 2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), 2021, pp. 735-738, doi: 10.1109/RTEICT52294.2021.9573964.

[4]   Dakopoulos, Dimitrios, and Nikolaos G. Bourbakis. "Wearable obstacle avoidance electronic travel aids for blind: a survey." IEEE Transactions on Systems, Man, and Cybernetics, Part C(Applications and Reviews) 40, no. 1 (2010): 25-35.

[5]   Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster R-CNN: Towards real-time object detection with region proposal networks." In Advances in neural information processing systems, pp. 91-99. 2015.

[6]   Dumitru Erhan, Christian Szegedy, Alexander Toshev, and Dragomir Anguelov Google, Inc. 1600 Amphitheatre Parkway, Mountain View (CA), 94043, USA.

[7]   Fares Jalled, ´ Moscow Institute of Physics & Technology, Department of Radio Engineering & Cybernetics Ilia Voronkov, Moscow Institute of Physics & Technology, Department of Radio Engineering & Cybernetics

[8]   Xiaozhi Chen∗ , Kaustav Kundu∗ , Yukun Zhu, Huimin Ma, Sanja Fidler and Raquel Urtasun.

[9]   Ross Girshick1 Jeff Donahue1,2 Trevor Darrell1,2 Jitendra Malik1 1UC Berkeley and 2 ICSI {rbg,jdonahue,trevor,malik} @eecs.berkeley.edu.

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⊙ (24*7 Support on Whatsapp)