International Journal of
Optical Innovations & Research
Volume 2, Issue 1, 2024
January—June
DOI (Journal): 10.37591/IJOIR
https://journals.stmjournals.com/ijoir
Review IJOIR

Innovative Eyewear for the Visually Impaired

Vaishnavi Bhapkar[1], Sankita Salvi[2], Jankar Tejaswini[3], Bandal Rutuja[4], K.S. Khamkar[5,*]

Abstract

Object detection systems are essential tools for identifying and locating objects within images or videos. When integrated into spectacles or wearable devices, these systems provide users with real-time information about objects present in their surroundings. This functionality serves diverse purposes, such as assisting visually impaired individuals in navigating their environment or offering augmented reality data to workers during tasks. Region-based Convolutional Neural Networks (RCNN) represent a prominent machine learning model used extensively for object detection. The RCNN model operates in two main stages: initially employing a convolutional neural network (CNN) to extract distinctive features from the input image. Subsequently, it applies a region proposal algorithm to pinpoint potential object locations within the image. These proposed regions are then processed through a second CNN, which classifies them into either objects or background. Effectively, the RCNN model has demonstrated its capability to detect a broad spectrum of objects across various types of images and videos. Its proficiency lies in leveraging deep learning techniques to accurately identify and categorize objects, making it a versatile tool for applications ranging from enhancing accessibility for the visually impaired to improving productivity through augmented reality in industrial settings.

Keywords: Moving Object Detection Systems, Machine Learning, Region-based CNN, Spectacles, Algorithm

INTRODUCTION

In today's advanced technological landscape, artificial intelligence (AI) is constantly evolving and being integrated into various technologies that enhance quality of life, safety, security, entertainment, and more. AI systems are designed to autonomously make decisions and learn, with the overarching goal of automating tasks across different domains. Object identification represents one such application of AI that significantly contributes to safety and quality of life improvements.

The advent of new media, particularly smart glasses, has revolutionized human behavior by bridging the gap between physical and digital realms. For individuals with visual impairments and blindness, who often face challenges in navigating their environments independently, innovative solutions like the Intelligent Spectacular for Blind People aim to address these difficulties. This wearable device employs artificial intelligence and computer vision techniques to assist blind and visually impaired (BVI) individuals in performing everyday tasks without relying on external assistance [1–5].

Communication within the Intelligent Spectacular for Blind People system is facilitated through WIFI and Bluetooth technologies. The device's camera module continuously captures real-time images of surroundings, which are then

***Author for Correspondence**
K.S. Khamkar
E-mail: vaishb2709@gmail.com

[1-4]Student, Department of Computer Engineering, RDTC's Shri Chhatrapati Shivajiraje College of Engineering, Dhangwadi, Bhor, Pune Maharashtra, India
[5]Professor, Department of Computer Engineering, RDTC's Shri Chhatrapati Shivajiraje College of Engineering, Dhangwadi, Bhor, Pune Maharashtra, India

Received Date: May 31, 2024
Accepted Date: June 27, 2024
Published Date: July 10, 2024

Citation: Vaishnavi Bhapkar, Sankita Salvi, Jankar Tejaswini, Bandal Rutuja, K.S. Khamkar. Innovative Eyewear for the Visually Impaired. International Journal of Optical Innovations and Research. 2024; 2(1): 25–31p.

transmitted for classification and identification. These images are processed in the cloud server, where they undergo enhancement and object detection procedures. Once objects are identified, labels and relevant information are added to the images and sent back to the device for user feedback.

Object recognition, a core computer vision technique, enables AI systems to identify objects in images or videos using deep learning and machine learning algorithms. The primary objective of object detection is to teach AI systems the innate human ability to effortlessly recognize various items in visual media, such as people, vehicles, animals, and objects. By training AI systems in this way, they gain the capability to analyze images, extract meaningful elements, and comprehend the contents depicted.

In essence, the integration of AI-driven object recognition and computer vision technologies in devices like the Intelligent Spectacular for Blind People exemplifies how advanced technology can empower individuals with disabilities, enhancing their autonomy and enabling greater participation in daily life activities [6].

RELATED WORK

Object detection methods using convolutional neural networks (CNNs) offer different trade-offs between speed and accuracy. SSD (Single Shot Multibox Detector) with MobileNetV1 provides fast detection speeds but sacrifices accuracy compared to Faster R-CNN with InceptionV2, which offers higher accuracy but is slower in detection speed. These findings highlight a clear trade-off where choosing between the models depends on the specific application requirements.

For applications where quick detection is crucial, especially in real-time scenarios, SSD with MobileNetV1 is recommended due to its rapid processing capabilities. On the other hand, Faster R-CNN with InceptionV2 is preferable when high accuracy in detection is paramount, despite its slower speed [7].

Looking ahead, both models will be employed as the vision system for a bomb disposal robot aimed at detecting improvised explosive devices (IEDs). This choice reflects a strategic decision to leverage SSD with MobileNetV1 for swift initial detection in dynamic and potentially hazardous environments. Meanwhile, Faster R-CNN with InceptionV2 will serve to ensure precise identification and classification of objects, crucial for the safety and effectiveness of the robot's operations.

In summary, the selection between SSD with MobileNetV1 and Faster R-CNN with InceptionV2 depends on balancing the need for speed versus accuracy, tailored to the specific demands of applications such as bomb disposal robotics. The Figure 1. Shows the low-resolution image and high-resolution image.

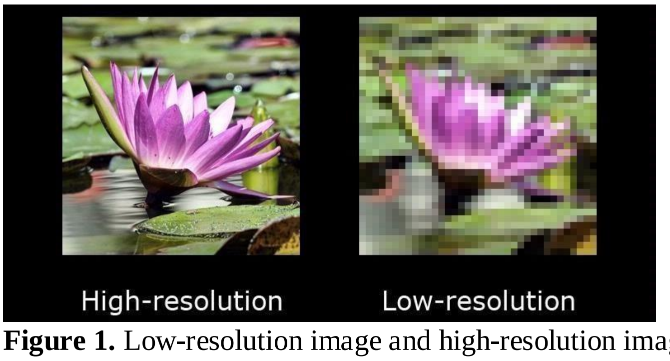

Figure 1. Low-resolution image and high-resolution image.

The proposed deep learning architecture for moving object detection in moving camera environments, like dashcam footage, integrates two specialized networks: one emphasizing appearance and the other mobility. This innovative approach effectively mitigates background interference, even with unrestricted camera movement, marking a notable advancement. Experimental validation against leading methods has confirmed its superior performance [8].

Moreover, the technique offers the distinct advantage of operating in real-time, processing up to 50 frames per second. This capability makes it well-suited for demanding real-world applications, including scenarios encountered in autonomous vehicles. Overall, this framework represents a significant step forward in enhancing moving object detection accuracy and efficiency in dynamic camera settings.

Moving Object Detection for Event-based Vision using Graph Spectral Clustering: GSCEventMOD is an innovative tool designed for detecting moving objects using event-based vision and graph spectral clustering. It has proven effective in challenging scenarios such as high-speed movements and sudden changes in lighting conditions, leveraging the capabilities of neuromorphic vision sensors. These sensors capture scene dynamics directly, reducing the need for extensive preprocessing of data.

In the realm of event-based vision, GSCEventMOD has demonstrated superior performance compared to traditional methods. Validation using both synthetic and real-world data across diverse environments underscores its adaptability and robustness.

Overall, GSCEventMOD is anticipated to find widespread application across various computer vision domains, including autonomous vehicles, robotics, and remote surveillance. Its ability to handle dynamic scenes efficiently positions it as a promising advancement in object detection and tracking technologies [9].

Moving Object Detection by a Mounted Moving Camera: Using the pyramidal Lucas-Kanade approach, locate and follow interest points in a series of video frames. Subsequently, it computes camera motion by assuming that the motion vectors with the highest frequency are associated with camera motion. The frame difference approach is used to detect moving objects after the motion of the camera has been eliminated. utilizing adaptive thresholding while accounting for various lighting scenarios within a single video. Moving Figure 2. Shows the object detection using Intelligent Spectaculars

Using frame differencing and the summing technique, moving object detection and segmentation This method is simple and low computationally complex in comparison to traditional object recognition and segmentation approaches. This method separates moving objects from the static backdrop fast and accurately. However, if they are greater than the cut-off number, they are also classified as moving objects because this technique does not consider the shadows of moving objects.

PROPOSED WORK

In this section, we provide a detailed overview of the camera motion estimation and moving object detection methods integrated into our system.

The global population includes a significant number of individuals who are blind. These individuals face challenges in perceiving and understanding their surroundings due to their inability to see. Our innovation directly addresses this issue through the “Intelligent Spectacular for Blind People.” This wearable device, resembling goggles, empowers blind individuals by enabling them to comprehend and identify objects through auditory feedback.

The system utilizes Computer Vision technology, specifically OpenCV, for capturing images and implementing preprocessing techniques. These techniques are essential for optimizing image data by focusing only on regions pertinent to object detection. This preprocessing step ensures that the subsequent object detection algorithms operate efficiently and accurately.

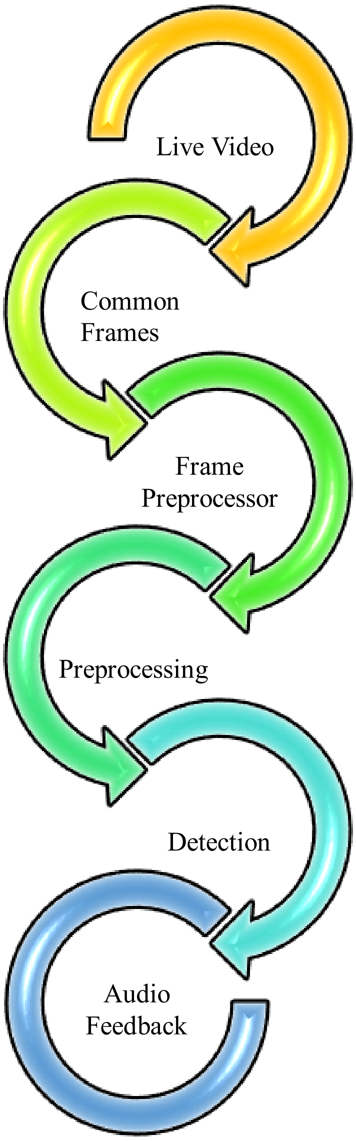

Figure 2. Moving Object Detection using Intelligent Spectaculars.

Overall, our invention aims to enhance the independence and awareness of blind individuals by converting continuous visual information into accessible auditory cues, thereby enabling them to navigate and interact with their environment more effectively.

Raspberry Pi 3

The Raspberry Pi Foundation, in partnership with Broadcom, developed a series of compact single-board computers called the Raspberry Pi. Initially, the Raspberry Pi initiative aimed to enhance computer science education in schools and underserved communities by providing affordable computing solutions.

The Raspberry Pi line of single-board computers, created by the UK-based Raspberry Pi Foundation, receives video input from cameras and executes primary algorithms. In the context of object detection, these devices deploy models designed to identify objects and deliver feedback to users.

In continuous video streams, the system employs background subtraction techniques to detect and categorize moving objects. Once an object is successfully identified, the system communicates this information to the user through audible feedback.

In essence, the Raspberry Pi serves as a versatile platform for implementing object detection models, utilizing video input to perform real-time analysis and provide actionable feedback to users via audio cues. This capability supports various applications, from educational projects to practical solutions in computer vision and automation.

Camera

The camera of our Intelligent Spectacle for Blind People is fixed to the spectacles' bridge (wearable goggles). The camera can capture excellent video images because it has 8.0 megapixels. The is inserted in glasses and connected to the device in a way that records a large portion of the user's environment. First, the camera takes a sequence of photos of continuously moving objects to create a collection of short video clips. An image is first labeled by RCNN. After predicting whether the image has one or more elements, it identifies the class. After utilizing a bounding box to locate the object, object localization is carried out in the image.

Moving Object Detection

Detecting moving objects can be challenging but leveraging RCNN (Region-based Convolutional Neural Network) with the Raspberry Pi 3 Model B+ significantly accelerates the process. RCNN is a deep learning technique tailored for image object detection. It segments images into regions and then classifies each region as containing an object or not, based on training with a large dataset of labeled images.

After training, the RCNN model can be deployed on the Raspberry Pi 3 Model B+, which offers sufficient processing power to run the model in real-time. This makes the Raspberry Pi 3 Model B+ well-suited for immediate processing as soon as the camera captures a frame, swiftly labeling bounding boxes around detected objects. These bounding boxes are instrumental in tracking the movement of objects over time.

To achieve accurate results, the system compares these detected objects against a pre-established dataset. This dataset consists of labeled images used during the RCNN model training, enabling the model to recognize objects and their distinguishing features. When presented with a new image, the RCNN model analyzes it using the bounding boxes to identify and classify objects based on their learned features from the dataset.

This interaction between the detected objects and the dataset is crucial for ensuring precise and consistent outcomes. Overall, combining RCNN with the Raspberry Pi 3 Model B+ forms a robust solution for effectively identifying moving objects, demonstrating its capability as a powerful tool in various applications requiring real-time object detection and tracking.

Headphones

In the previously described system, the results of object detection are conveyed to the user through audio feedback delivered via headphones. This setup allows users to receive real-time updates on detected objects without needing to visually inspect a screen. Depending on user preferences and needs, the audio feedback system can be tailored to provide various types of information based on the detected objects. This information serves to assist users in making informed decisions or taking appropriate actions.

The use of audio feedback is particularly advantageous in scenarios where users must maintain constant awareness of their surroundings, such as in security or surveillance applications. It enables users to stay vigilant and responsive to changes in their environment without the continuous need to visually monitor a display.

Overall, combining object detection with audio feedback represents a potent tool with diverse applications. It enables real-time detection and analysis of objects, offering users the critical information necessary for making informed decisions. In the context of the Intelligent Spectacular for Blind People described earlier, the camera module continuously captures real-time images of objects, which are then transmitted for categorization and recognition on a Raspberry Pi 3-based model. Following image enhancement and object identification, the labeled information is sent back to the user through auditory cues upon detecting objects in motion.

Button: Header pins are used to connect the button to the Raspberry Pi. It is integrated, so the blind person can press it in the improbable event of a tragic catastrophe. Notification that the blind person is in danger and the blind person's current location are sent to the person whose phone number has been saved.

Buzzer: A sound pulse in the ultrasonic range is produced by an ultrasonic sensor. The ultrasonic sensor helps blind people navigate their surroundings by sounding a buzzer to warn them of nearby items that could be barriers.

CONCLUSION

We explore a method for detecting moving objects in images, acknowledging that traditional techniques assume each positive image features only one object. Our study explores image enhancement, motion detection, object tracking, and behavior analysis to extract comprehensive information from images.

One limitation we address is the inefficiency of temporal differencing, which fails to capture all significant foreground pixels, particularly when objects have uniform textures or move slowly. Additionally, when a moving foreground object halts, temporal differencing cannot detect changes between consecutive frames, resulting in the loss of object tracking.

This paper offers valuable insights into these critical challenges and advocates for further research in the domains of moving object identification and computer vision. In our investigation, various estimation methods are applied in kernel tracking strategies to accurately delineate the target object's region. Currently, Mean-shift tracking and particle filtering stand as widely recognized and recommended approaches in kernel tracking.

Contour tracking within kernel tracking methods can be categorized into two approaches based on how contours evolve: the state space technique and the energy function minimization approach. The state space technique models object contours by predicting their transitions over time, employing dynamic models for future position estimation. In contrast, the energy function minimization approach optimizes an energy function that quantifies discrepancies between anticipated and observed contours. Techniques like active contours (snakes) fall into this category, adjusting their shape to conform to object boundaries based on energy minimization principles.

In summary, this study contributes crucial insights into the realm of moving object detection, emphasizing the importance of advancing techniques in computer vision. By exploring diverse tracking and estimation methodologies, we aim to enhance the accuracy and reliability of identifying and tracking moving objects in complex visual scenarios.

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 D, Bourbakis NG. Wearable obstacle avoidance electronic travel aids for blind: a survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews). 2009 Jul 14;40(1):25-35
5. Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems. 2015;28.
6. Erhan D, Szegedy C, Toshev A, Anguelov D. Scalable object detection using deep neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition 2014 (pp. 2147-2154).Jalled F, Voronkov I. Object detection using image processing. arXiv preprint arXiv:1611.07791. 2016 Nov 23.
7. Chen X, Kundu K, Zhu Y, Ma H, Fidler S, Urtasun R. 3d object proposals using stereo imagery for accurate object class detection. IEEE transactions on pattern analysis and machine intelligence. 2017 May 19;40(5):1259-72.
8. Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition 2014 (pp. 580-587).
9. Roopa GM, Prakash C, Pradeep N. Computer Vision-Based Assistive Technology for Blind and Visually Impaired People: A Deep Learning Approach. Computer Assistive Technologies for Physically and Cognitively Challenged Users. 2023 Mar 22:48.

