Smart Glasses with Blind Assistance System: Real-Time Object Detection with Voice Alerts

Guide: Prof.Gunaware.N
Jadhav Soham Ganesh, **Dange Asmita Chandrakant**, **Walke Shivam Ramesh**, **Bhujbal Bhavesh Gangadhar**

[12345] *Department of Computer Engineering, HSBPVT'S FOE, Maharashtra, India*

ABSTRACT

The "Smart Glasses with Blind Assistance System: Real-Time Object Detection with Voice Alerts" is an innovative technology designed to empower individuals with visual impairments. This system, developed using Python, combines real-time object detection and audio feedback to assist users in navigating their surroundings independently and safely. Beyond navigation, it can also function as an alarm alert system, providing auditory cues upon detecting objects. The technology's potential extends to becoming a hardware product for biomedical applications. Ongoing research aims to enhance accuracy through model retraining and implementation distance calculations for detected objects. Additionally, plans include deploying the system on hardware devices like smart glasses and expanding accessibility through a dedicated mobile app. This smart glasses-based solution holds great promise in improving the quality of life and independence for visually impaired individuals.

Keywords: - *Internet of Things (IoT), Object detection, YOLO V3, Voice feedback, raspberry PI 4b, wearable device*

1. INTRODUCTION

The "Smart Glasses with Blind Assistance System: Real-Time Object Detection with Voice Alerts" represents a groundbreaking innovation in assistive technology, specifically tailored for individuals with visual impairments. Developed at the intersection of cutting-edge computer vision and audio feedback systems, these smart glasses offer real-time object detection capabilities. By leveraging advanced algorithms, the system can identify objects and provide instant voice alerts, enabling users to navigate their environment with newfound confidence and independence. This introduction explores the fusion of wearable technology and artificial intelligence, highlighting the potential to revolutionize the daily lives of individuals living with visual challenges.

Key Features of Smart Glasses with Blind Assistance System: Real-Time Object Detection with Voice Alerts:

1. Real-Time Object Detection: The system provides instantaneous detection of objects in the user's surroundings, offering valuable awareness and safety.
2. Voice Alerts: Users receive auditory alerts that describe detected objects, aiding in navigation and enhancing situational awareness.
3. Wearable Design: The technology is integrated into a pair of smart glasses, ensuring a discreet and hands-free user experience.
4. Customizable Alerts: Users can customize the type and frequency of voice alerts to match their preferences and specific needs.
5. Obstacle Avoidance: The system assists users in identifying and avoiding obstacles, such as curbs, stairs, and potential hazards.

Components of Smart Glasses with Blind Assistance System:

1. Camera: A built-in camera captures the user's surroundings, providing visual input for object detection.
2. Object Detection Software: Advanced computer vision algorithms analyze the camera feed in real time to identify objects and obstacles.
3. Voice Feedback System: An audio system delivers voice alerts and descriptions to the user through bone conduction or earphones.
4. Battery: The glasses are equipped with a rechargeable battery to power the device throughout the day.
5. Raspberry Pi 4B: A powerful onboard processor that handles real-time object detection and voice alert generation.

2. LITERATURE REVIEW

[1]These key features and components collectively create a comprehensive and user-friendly Smart Glasses with Blind Assistance System, enhancing the mobility, independence, and safety of individuals with visual impairments.

[2]Several existing technologies contribute to the development of Smart Glasses with Blind Assistance Systems, focusing on real-time object detection with voice alerts. Computer vision algorithms, particularly those employing convolutional neural networks (CNNs), enable sophisticated object recognition capabilities. Deep learning frameworks such as YOLO (You Only Look Once) and SSD (Single Shot Multi-Box Detector) have gained prominence for their accuracy and real-time processing, making them instrumental in object detection tasks.

[3]Additionally, advancements in natural language processing (NLP) and speech synthesis technologies facilitate the generation of clear and contextually relevant voice alerts. Wearable sensors like accelerometers and gyroscopes enhance the system's spatial awareness, enabling a more comprehensive understanding of the user's environment. Furthermore, developments in bone conduction technology and miniature audio components have improved the delivery of voice alerts without obstructing ambient sounds, ensuring a seamless user experience.

[4]Cloud computing and edge computing solutions have also contributed significantly, enabling real-time data processing and enhancing the system's responsiveness. Furthermore, augmented reality (AR) technologies have been integrated into some smart glasses, offering potential applications for visually impaired users, including real-time annotations and additional auditory cues for object recognition.

[5]These existing technologies collectively form the foundation for the Smart Glasses with Blind Assistance System, driving innovation in the realm of assistive devices for individuals with visual impairments.

3. PROPOSED SYSTEM

The proposed system is a Smart Glasses with Blind Assistance System that combine’s real-time object detection and voice alerts. This technology empowers individuals with visual impairments by providing instant auditory cues about their surroundings. The system is designed as a wearable solution integrated into a pair of smart glasses. It offers features such as obstacle avoidance, facial recognition, text reading, color recognition, and customizable voice alerts. The system's components include a camera, object detection software, voice feedback system, microphone, processor, display (optional), and wireless connectivity. The proposed system aims to enhance the independence, safety, and overall quality of life for individuals with visual impairments.

3.1. Hardware Components used:

1.Laptop or PC

2.Raspberry Pi 4B

3.720p HD Camera

4.Wearable Glasses

5.DC power bank

Development environment with necessary software tools (IDEs, code editors).

1. Laptop or PC –

Laptop with an i5 processor minimum 8 GB RAM keyboard and Mouse with the help of this we developed the initial program for object detection

2**. Raspberry Pi 4B-**

A Raspberry Pi 4B is a key component in object detection as it plays a central role in detecting objects and giving the voice feedback process.

Fig.1

Brand	Raspberry Pi
Model	Raspberry Pi Model 4B
RAM	4GB
Processor	Broadcom 2711, 64-bit quad-core Cortex-A72 processor
USB Ports	2 x USB 3.0
Pins	40 GPIO Pins
Bluetooth	Bluetooth 5.0
Dimension	85 x 56 x 19(mm)
Weight	65 grams

3. 720p HD Camera-

An "HD Ready" camera refers to a camera capable of capturing high-definition (HD) video or images, typically at resolutions of 720p, providing good-quality visuals for various purposes, such as photography, videography, and online content creation.

Fig.2

- 720p 30fps
- UVC, plug-and-play
- 60° wide angle view
- Up-down adjustable
- Swivel 270°adjustable
- Adjustable clip

4. **Wearable Glasses** –

We use wearable glasses with a camera attached to them for the user to wear them comfortably and they can use it for their convenience

Fig.3

5. **DC power bank-**

A DC power bank is a portable device that stores electrical energy and provides direct current (DC) output to charge Raspberry PI

A battery pack with a USB to type C cable can also power a Raspberry Pi. This battery pack provides up to 3.1A of current at 5V. Connect the Raspberry Pi to your battery pack.

Fig.4

3.2 Software Requirements:

To work with YOLO (You Only Look Once) version 3 for object detection, you'll need a combination of software tools and libraries. Here is a list of software requirements for YOLO v3 object detection:

1. Deep Learning Framework:
   - YOLO v3 is typically implemented using deep learning frameworks such as:
     - Darknet: YOLO's official framework, which is written in C and CUDA for GPU acceleration.
     - PyTorch: YOLOv3 implementations are available in PyTorch for flexibility and ease of use.
     - Tensor Flow: You can find YOLOv3 implementations in Tensor Flow as well.

2. CUDA and cuDNN:
   - If you plan to use GPU acceleration for training and inference, you'll need NVIDIA's CUDA and cuDNN libraries.

3. Python:
   - Python is commonly used for running scripts, data preparation, and working with deep learning frameworks.

4. OpenCV:
   - OpenCV (Open Source Computer Vision Library) is a valuable tool for image and video processing, which is often used in conjunction with YOLO for real-time object detection.

5. Labeling Tools:
   - Labeling tools such as Labeling, VoTT (Visual Object Tagging Tool), or Labelbox are essential for creating annotated datasets required for training the YOLO model.

6. Annotated Dataset:
   - A dataset of labeled images and corresponding annotations is needed for training your YOLO v3 model. This dataset should include images with bounding boxes and class labels.

7. Pretrained YOLO Weights:
   - Pretrained YOLO v3 weights are available, which can be used as a starting point for transfer learning or fine-tuning on your specific dataset.

8. Training Hardware:
   - Depending on the scale of your project, you might need a computer with a capable GPU for training. YOLO v3 benefits significantly from GPU acceleration.

9. Operating System:
   - YOLO can be used on various operating systems, including Linux, Windows, and macOS. The choice depends on your preference and hardware compatibility.

3.3 Algorithm Used:

YOLO V3-

YOLO is one of the most popular object detection models used for real-time object detection and classification in Computer Vision.
It uses CNN layers called Darknet-53.
It was originally developed by Joseph Redmon, Ali Farhadi, and Santosh Divvala.
YOLO achieves the highest accuracy in real-time speed.
It belongs to a stage detector that processes the input in one forward pass of the Convolution Neural Network.
After many two-stage models like R-CNN, FAST R-CNN, etc. YOLO is the model which changed the field of object detection in terms of real-time.
The Yolo model has had many upgrades through the years. The timeline of upgraded model releases can be seen in the below figure**.**

4.0 CONCLUSION & FUTURE WORK-

Intelligent object tracking and audio feedback have been developed in Python to facilitate real-time navigation for individuals with visual impairments. This model not only aids in navigation but can also serve as an alarm alert system by providing audio alerts upon object detection. To enhance its accessibility, there is potential to transform it into a hardware product for application in the biomedical field, where it can assist visually impaired individuals in navigating various environments. Future research and development efforts should prioritize model retraining to improve accuracy. While the current model achieves 80% accuracy, fine-tuning with real-world data can push it beyond 90%. Furthermore, deployment on hardware devices like smart glasses can significantly enhance the user experience and navigation capabilities for individuals with visual impairments. Calculating the distance to detected objects is a feature to consider for future versions. Finally, the project will be extended to a mobile app, broadening its user base and accessibility.

5. REFERENCES –

- https://medium.com/beingryaan/real-time-object-detection-along-with-distance-and-voice-alerts-for-blinds-a-blind-assistance-1708b97c3ecc
- https://www.mdpi.com/2079-9292/10/22/2756
- https://ieeexplore.ieee.org/abstract/document/8740604
- https://ieeexplore.ieee.org/abstract/document/8941141
- https://ieeexplore.ieee.org/abstract/document/9544598
-
- Gai, Yakun Liu, Jing Zhang & Gang Jing (2021) An improved Tiny YOLOv3 for real-time object detection, Systems Science & Control Engineering, 9:1, 314-321, DOI: 10.1080/21642583.2021.1901156
- 2) Guo, Shih-Sian & Lee, Kuo-Hua & Chang, Liyun & Tseng, Chin-Dar & Sie, Sin-Jhe & Lin, Guang-Zhi & Chen, Jih-Yi & Yeh, Yi-Hsin & Huang, Yu-Jie & Lee, Tsair-Fwu. (2022). Development of an Automated Body Temperature Detection Platform for Face Recognition in Cattle with YOLO V3-Tiny Deep Learning and Infrared Thermal Imaging. Applied Sciences. 12. 4036. 10.3390/app12084036.
- 3) K. M and K. P. R, "Comparative Analysis of YOLOv3-320 and YOLOv3-tiny for the Optimised Real-Time Object Detection System," 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2022, pp. 495-500, doi: 10.1109/ICIEM54221.2022.9853186.
- 4) Taheri Tajar, A., Ramazani, A. & Mansoorizadeh, M. A lightweight Tiny-YOLOv3 vehicle detection approach. *J Real-Time Image Proc* **18**, 2389–2401 (2021). https://doi.org/10.1007/s11554-021-01131-w
- 5) C. Tan, D. CHEN, H. Huang, Q. Yang and X. Huang, "A Lightweight Underwater Object Detection Model: FL-YOLOV3-TINY," 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2021, pp. 0127-0133, doi: 10.1109/IEMCON53756.2021.9623066.

- 6) Yu, Z., Bouganis, CS. (2020). A Parameterisable FPGA-Tailored Architecture for YOLOv3-Tiny. In: Rincón, F., Barba, J., So, H., Diniz, P., Caba, J. (eds) Applied Reconfigurable Computing. Architectures, Tools, and Applications. ARC 2020. Lecture Notes in Computer Science(), vol 12083. Springer, Cham. https://doi.org/10.1007/978-3-030-44534-8_25
- 7) Adarsh, Pranav & Rathi, Pratibha & Kumar, Manoj. (2020). YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. 687-694. 10.1109/ICACCS48705.2020.9074315.
- 8) Sujee, R., Shanthosh, D., Sudharsun, L. (2020). Fabric Defect Detection Using YOLOv2 and YOLO v3 Tiny. In: Chandrabose, A., Furbach, U., Ghosh, A., Kumar M., A. (eds) Computational Intelligence in Data Science. ICCIDS 2020. IFIP Advances in Information and Communication Technology, vol 578. Springer, Cham. https://doi.org/10.1007/978-3-030-63467-4_15
- 9) Chen, Xiaojuan and Zhengyang Yu. "An improved pedestrian detection algorithm based on YOLOv3." *International Conference on Electronic Information Engineering and Computer Science* (2023).
- 10) A. Sumarudin, Adi Suheryadi, Alifiah Puspaningrum, Aditya Rifqy Fauza; Implementation YOLOv3 for symptoms of disease in shallots crop. *AIP Conference Proceedings* 27 April 2023; 2646 (1): 020023. https://doi.org/10.1063/5.0113748
- 11) Redmon, Joseph, et al. "You Only Look Once: Unified, Real-Time Object Detection." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2016, pp. 779–88, doi:10.1109/CVPR.2016.91."
- 12) Joseph Redmon, Ali Farhadi "YOLOv3: An Incremental Improvement" 2018 https://doi.org/10.48550/arXiv.1804.02767
- 13) Baokai Liu, Fengjie He, Shiqiang Du, Jiacheng Li, Wenjie Liu "An advanced YOLOv3 method for small object detection" December 2022, https://doi.org/10.48550/arXiv.2212.02809