Literature Survey on Laptop-Assisted Smart Glasses for Text Reading, Object Detection and Obstacle Warning for Blind and Visually Impaired People

A Review Supporting the NETRA Project

Prathamesh Kaklij#1, Sheryl D'Souza#2, Shreyas Jadhav#3, Komal Bhange#4

Project Guide: Prof. Tejal Sonawane#5

#Department of Computer Engineering, Guru Gobind Singh Polytechnic, Nashik, Maharashtra, India

1. sheryldsouza770@gmail.com , 2. prathameshkaklij212007@gmail.com, 3. bhangekomal59@gmail.com,

***Abstract*—**This paper presents a literature survey prepared for NETRA, a laptop-assisted pair of smart glasses planned for blind and visually impaired people. The survey reviews published work on the three functions that the project combines: reading printed English text, recognizing everyday objects and warning about nearby obstacles. Earlier studies on camera-based travel aids, optical character recognition, real-time object detection, ultrasonic distance sensing and wearable assistive devices are examined and compared. The comparison shows that many existing systems cover only one function, that obstacle warning often depends on a single front-facing sensor, and that little is said about how spoken output should be prioritized. These observations point to the need for an affordable wearable system that joins text reading, object detection and direction-wise obstacle warning and announces danger before other information. The findings form the basis for the design decisions of NETRA.

***Keywords*—**assistive technology; smart glasses; optical character recognition; object detection; ultrasonic obstacle warning; text-to-speech

I. INTRODUCTION

Vision loss touches almost every activity of daily living, so technology that takes over even a few visual tasks can make a real difference to a person's independence [1]. Computer vision seems a natural fit for this purpose because it can stand in for the lost sense of sight. Even so, Manduchi and Coughlan observe that very few computer vision systems are actually used by visually impaired people [1].

This survey is organized around the three functions that NETRA combines: reading printed text, identifying objects and warning about obstacles. NETRA is a low-cost smart-glasses prototype. The glasses carry a USB camera, distance sensors, push buttons and an Arduino Nano or ESP32 controller, while a laptop runs OpenCV, English optical character recognition and an object-detection model. The result is converted into short, prioritized speech and played through an open-ear earphone, so the user can still hear the surroundings.

The survey draws on journal papers, conference papers, a book chapter, one preprint and one university project report. The preprint and the project report have not gone through the same review as the other sources; they are marked as such in the reference list and are used only to illustrate design choices. Section II reviews related work in five groups, Section III compares the systems, Section IV states the research gap and the way NETRA responds to it, and Section V concludes.

II. REVIEW OF RELATED WORK

A. Computer Vision and Travel Aids

Manduchi and Coughlan [1] reviewed the areas of assistive technology that have already been developed for visually impaired people and the part that computer vision can play in them. They also analyzed why earlier attempts failed and suggested promising research directions. Their description of electronic travel aids, which typically rely on range sensors such as sonar, active triangulation and stereo vision,

is a useful starting point for the choice of sensors in NETRA.

Budrionis et al. [2] followed the PRISMA method to review smartphone-based travel aids. They found that few studies combined haptic interfaces with computer vision, that modern deep-network methods were rarely used, and that the prototypes built by researchers did not match well with the needs expressed by blind expert users. This mismatch is a reminder that new prototypes should be checked against real user needs.

Fei et al. [3] reviewed travel aids based on machine vision and compared them with aids that use other sensing techniques. They grouped the functions from low-level image processing, such as finding road regions and obstacles, to high-level functions such as recognizing digital tags and text. NETRA covers this range by using the camera for text and objects and separate distance sensors for obstacles.

B. Text Reading Using OCR

Reading printed text depends on a dependable OCR engine. Smith [4] gave a detailed overview of Tesseract, an engine that was developed at HP between 1984 and 1994 and later released as open source. The overview highlights its line-finding step, its feature and classification methods and its adaptive classifier. Because it is open source, Tesseract suits a low-cost prototype that must read printed English.

Before text reaches the OCR engine, the image usually needs clean-up such as resizing, grayscale conversion and thresholding. OpenCV [5], an open-source library for image processing and computer vision, offers these operations and is the library selected for the laptop pipeline of NETRA.

C. Object Detection

Real-time object recognition became practical with single-network detectors. Redmon et al. [6] introduced YOLO, which treats detection as a regression problem and predicts bounding boxes and class probabilities directly from the whole image in one evaluation. Their base model processed 45 frames per second, and a smaller version, Fast YOLO, processed 155 frames per second. They also reported that YOLO makes more localization errors than some competing systems but is less likely to report false detections on the background.

For NETRA, this speed suits a live camera view handled by a laptop, and the low rate of background false alarms matters when the user depends on short spoken messages. The final model will be chosen during testing, but detectors of this type are strong candidates.

D. Obstacle Detection Using Distance Sensors

Yazid and Othman [7] built a wearable cap with an ESP32 controller, an ultrasonic sensor and a buzzer. The device detected obstacles between 2 cm and 300 cm and used three sound patterns for different distance ranges: a continuous beep up to 100 cm, a long beep from 100 to 200 cm and a short beep from 200 to 300 cm. A GPS module sent location alerts through Telegram. The authors reported a practical limitation: when the cap was tilted 15 degrees downward, an obstacle 300 cm away fell outside the sensor beam and went undetected. The sensor also showed a minimum reading of 6 cm because of its dead zone. They further noted that a white cane cannot identify obstacles above waist level.

Ali et al. [8] used an Arduino Uno with an HC -SR04 ultrasonic sensor and gave sound feedback to the user. The sensor was placed at eye level to guard against head collisions. By the authors' own account, the model does not tell the type, position, shape or size of the obstacle, detects only objects about one to five meters ahead, and its loudness-based alert is not suitable for users with hearing problems. This work is a preprint and has not been peer reviewed.

Both studies show that ultrasonic sensing is cheap and simple, but a single front sensor and a buzzer give limited information. This supports the plan in

NETRA to measure distance on the left, centre and right and to announce the direction in words.

E. Integrated Smart Glasses and Wearable Systems

Srivastava and Singh [9] proposed smart hand gloves that combine obstacle detection, object identification and an OCR-based text-to-speech converter. As noted by Yazid and Othman [7], frequent hand movement can reduce the accuracy of object detection in such a design. Placing the camera on glasses keeps it aligned with the head direction of the user.

Aljamil [10] described smart glasses for blind and visually impaired people built around a Raspberry Pi 4. The system uses OpenCV and OCR to identify text, a Single Shot Multibox Detector to find objects, text-to-speech for audio output and ultrasonic sensors for hazard alerts. This project report shows that the three functions can be placed in one pair of glasses, although all processing is done on the small on-board computer.

Kumari and Hammady [11] proposed a framework on Vuzix Blade 2 smart glasses that joins object detection, distance estimation and OCR with audio feedback. They compared YOLOv11 and YOLOv12 for accuracy and computational efficiency and reported that the system could detect obstacles within one meter, estimate distances and convert printed text into speech in real time. Their design is tied to one specific pair of commercial glasses.

III. COMPARATIVE SUMMARY

Table I compares the system-level studies reviewed above with the proposed NETRA prototype. The review papers [1]–[3] and the component studies [4]–[6] are not included because they describe methods, not complete devices.

TABLE I. COMPARISON OF REVIEWED SYSTEMS

Ref.	System and main techniques	Text	Object	Obstacle	Limitation noted
[7]	Wearable cap; ESP32, ultrasonic sensor, buzzer, GPS alerts	No	No	Yes	Downward-tilted beam missed a distant obstacle; sound patterns give distance only
[8]	Arduino Uno, HC-SR04 sensor, sound alert (preprint)	No	No	Yes	Does not give obstacle type, position or size; limited to about 1 to 5 m ahead
[9]	Smart hand gloves; obstacle detection, object identification, OCR with speech	Yes	Yes	Yes	Hand movement may reduce detection accuracy, as noted in [7]
[10]	Smart glasses; Raspberry Pi 4, OpenCV, OCR, SSD detector, ultrasonic sensors, speech (project report)	Yes	Yes	Yes	Not peer reviewed; processing done on a small on-board computer
[11]	Vuzix Blade 2 glasses; YOLOv11 and YOLOv12, distance estimation, OCR, audio	Yes	Yes	Yes	Built around one commercial pair of glasses; obstacles tested within one meter

Ref.	System and main techniques	Text	Object	Obstacle	Limitation noted
NETRA	Laptop-assisted glasses; USB camera, distance sensors, OpenCV, OCR, object detection, speech	Yes	Yes	Yes	Prototype stage; indoor testing first; depends on a laptop

IV. RESEARCH GAP AND RELEVANCE TO NETRA

The reviewed work shows that each of the three functions is well studied on its own, and that a few systems [9]–[11] already bring them together. Within the material reviewed here, the following gaps remain, and each one shapes a design decision of NETRA.

- **Single-function devices.** Obstacle-only devices [7], [8] warn the user but cannot read a sign or name an object. NETRA adds text reading and object detection to the same glasses.
- **Direction of the obstacle.** The devices in [7] and [8] use one front-facing sensor. NETRA is planned with left, centre and right distance sensors so that the spoken warning can name the side of the obstacle.
- **Sensor placement.** The blind spot reported in [7] shows that mounting angle can hide an obstacle from the sensor. The mounting position and tilt on the NETRA frame will be tested before final integration.
- **Form of the alert.** Buzzer tones carry little meaning, and loudness-based alerts can exclude users with hearing difficulty [8]. NETRA uses short English speech, so the user receives words instead of tones.
- **Order of spoken messages.** The descriptions of the combined systems [9]–[11] reviewed here do not center on a rule that places danger ahead of other information. NETRA applies a priority order in which a very close obstacle is announced first, followed by a general warning, requested text, object results and navigation information.
- **Open-ear audio.** Because the user relies on hearing the surroundings, NETRA delivers speech through an open-ear earphone.
- **Cost and processing.** Compact on-board computers [10] or dedicated smart glasses [11] fix the hardware choice. NETRA moves heavy processing to a laptop so that the glasses stay simple and low in cost, at the price of depending on that laptop. The review in [2] stresses affordability for low-income settings, which supports this approach.

Following the lesson of [2], the prototype will be checked against user needs. NETRA is an assistive prototype only and is not a replacement for a white cane, a guide dog or mobility training.

V. CONCLUSION

This survey examined published work on camera-based travel aids, OCR, object detection, ultrasonic obstacle sensing and wearable assistive systems. The studies show that each function is mature enough for a student prototype, but that many devices offer only one function, use a single front sensor and give little attention to the order of spoken output. NETRA addresses these points by combining text reading, object detection and left, centre and right obstacle warning in one pair of laptop-assisted glasses, with short, prioritized speech through an open-ear earphone.

ACKNOWLEDGMENT

The authors thank Prof. Tejal Sonawane, Project Guide, Prof. Gayatri Jagtap, Head of the Department of Computer Engineering, and Guru Gobind Singh Polytechnic, Nashik, for their guidance and support.

REFERENCES

[1] R. Manduchi and J. M. Coughlan, “(Computer) vision without sight,” Commun. ACM, vol. 55, no. 1, pp. 96–104, Jan. 2012.

[2] A. Budrionis, D. Plikynas, P. Daniušis, and A. Indriulionis, “Smartphone-based computer vision travelling aids for blind and visually impaired individuals: A systematic review,” Assist. Technol., vol. 34, no. 2, pp. 178–194, 2022, doi: 10.1080/10400435.2020.1743381.

[3] Z. Fei, E. Yang, H. Hu, and H. Zhou, “Review of machine vision-based electronic travel aids,” in Proc. 23rd Int. Conf. Autom. Comput. (ICAC), 2017, doi: 10.23919/IConAC.2017.8082021.

[4] R. Smith, “An overview of the Tesseract OCR engine,” in Proc. 9th Int. Conf. Document Anal. Recognit. (ICDAR), Curitiba, Brazil, 2007, pp. 629–633.

[5] G. Bradski, “The OpenCV library,” Dr. Dobb’s J. Softw. Tools, 2000.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[7] M. I. Yazid and M. Othman, “Wearable obstacle detection for visual impairment people with GPS location,” Evol. Electr. Electron. Eng., vol. 5, no. 2, pp. 108–115, 2024, doi: 10.30880/eeee.2024.05.02.013.

[8] A. Ali, R. Qian, H. Akbar, and Z. Sartaj, “Electronic aid for unfamiliar environments,” Preprints, May 2024, doi: 10.20944/preprints202404.1630.v2 (preprint, not peer reviewed).

[9] N. K. Srivastava and S. Singh, “Netra: Smart hand gloves comprises obstacle detection, object identification & OCR text to speech converter for blinds,” in Proc. 5th IEEE Uttar Pradesh Sect. Int. Conf. Electr., Electron. Comput. Eng. (UPCON), 2018, pp. 1–4, doi: 10.1109/UPCON.2018.8596848.

[10] I. Aljamil, “Smart glasses for the blind and visually impaired (BVI),” project report, Newcastle Univ., Aug. 2024 (not peer reviewed).

[11] P. Kumari and R. Hammady, “A novel AI -driven assistive framework for enhanced navigation and text recognition for visually impaired using smart glasses,” in XR and Metaverse, Springer Proc. Bus. Econ., Springer, 2026, pp. 471–488, doi: 10.1007/978-3 -032-11983-4_35.