IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 1, June 2022

V-Eye : A Vision-based Navigation system for the Visually Impaired

Dr. G. Nanthakumar[1], Yogeshwaran. R[2], Santhosh. S[3], Pragadeeshwaran. R[4]

Assistant Professor & HoD, Department of Science and Computer Engineering[1]

Students, Department of Science and Computer Engineering[2,3,4]

Anjalai Ammal Mahalingam Engineering College, Kovilvenni, Tiruvarur, Tamil Nadu, India

Abstract: *Vision is one of the most vital senses in the human body, and it is essential for perceiving the world. However, millions of individuals around the world suffer from vision loss. They are having difficulty navigating their daily lives because they are unable to detect impediments in their environment, and one of their major challenges is recognising people. Other than automation, object detection is used in a variety of applications that have yet to be fully explored. This project includes one such application that employs detection to assist visually impaired individuals in identifying items ahead of them for safe navigation, as well as a face recognition system with aural output that can assist visually impaired persons in recognising faces. Speakers would provide them with voice-based assistance. We used a deep learning-based Faster Region-Convolutional Neural Network (Faster R-CNN) to detect and recognise humans and objects in the environment in this study. The Faster Region Convolution Neural Network technique processes and classifies the image captured by the camera. The audio jockey receives the detected image as an audio input. As a result, this model aids visually impaired persons in a more comfortable manner than white canes.*

Keywords: R-CNN :Region based Neural Network.

I. INTRODUCTION

The process of imparting data, information, and human intelligence to machines is known as artificial intelligence, or AI. Artificial Intelligence's major goal is to create self-contained devices that can think and act like people. Artificial neural networks, which enable machines to make judgments, are at the heart of deep learning. In some cases, deep learning-trained machines can now recognise images better than humans.

To create a smart glass system for those with vision impairments that uses computer vision techniques and deep learning models. Using real-time audio output, offer users with information about nearby items. Recognize relatives' and friends' faces and identity information.

.Vision is one of the most vital human senses, and it is essential for interpreting the world around us. However, millions of individuals around the world suffer from vision loss. They are having difficulty navigating their environment on a regular basis. Deep learning-based Faster Region -Convolutional Neural Network (Faster R -CNN) is used in this study to detect and distinguish humans and objects in the environment.

II. LITERATURE REVIEW

Sensor-based or vision-based approaches have been used in a lot of research on giving navigation help to visually impaired walkers, which will be covered below. In addition, Besides contrasting the planned global positioning system, In this section, we'll talk about vision-based localization and how it works. future advancements

A. Navigational Aids for the Blind and Visually Impaired

The most widely used sensor-based approach for outdoor applications. GPS [20] is used for navigation, however it is not precise enough. the current study's goals, particularly in terms of indoor research environments. A TP3 system, for example, may be used to support The ability of the visually impaired to navigate was discovered to be crucial. hampered by a scarcity of data for both indoor and outdoor applications[50] Environments Another new solution [8, 38] makes use of In conjunction with a smartphone app, beacons are used. Second, if the distances between beacons are too low, signal collision might occur, resulting in inaccurate position information. Finally, this method is difficult to implement.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 1, June 2022

Orientation information isn't available, and dynamic impediments aren't detected at all. As a result, following vision-based research has been conducted. emphasised the use of infrared-based depth cameras [27, 29] to Pedestrians should be able to navigate both indoors and outdoors with the help of impairments in vision While such technology is valuable up to a point, it isn't without its drawbacks. Experiments have revealed that these systems fail on a regular basis. infrared sensors, particularly when used outdoors, and that the systems are secure They are so hefty that consumers find it difficult to carry them. Context-aware navigation is a more accessible technique.

Technology that uses visual cues to guide vision impaired people along a path The use of a smartphone camera to recognise QR codes [25].Furthermore, none of the Above-mentioned systems can detect unanticipated obstructions, or give up-to-date orientation, or work equally well without it both indoors and out.The potential of a single person has received little attention. To improve, a monocular camera with computer vision is used. the efficiency with which such systems operate. In recent years, semantics has gained popularity. image segmentation [3, 24, 34] is a technique for deciphering images. By assigning labels to image pixels, the surrounding surroundings has been enhanced. become an important part of computer vision Recent research that used deep-learning networks [3, 24, 34] The superiority of such vision - based design is demonstrated by a architectures. Sensor-based approaches are preferred above others. Following serious consideration, As a result of this earlier work, we decided on a vision-based technique using a single camera to create a functional navigation system both internally and outside.

B. Localization Systems Based on Vision

There are two types of vision-based positioning technologies: visual simultaneous localization and visual localization. image-based localization and SLAM mapping SLAM for the eyes[17, 26, 28, 31, 33, 40, 44, 47], which is interactive and runs in real time. Constructs a 3D map and a 2D map at the same time, which is commonly used in robotics. calculates camera angles with relation to the camera's starting point However, It is hampered by the accumulation of errors, as well as its incapacity to recover. tracking issues, particularly when the camera rotates or moves. Its camera pose predictions are relative because it travels quickly.as opposed to absolute. [7, 9, 22, 23, 30, 31] Image-based localization On the other hand, [37, 43, 51] arrives to the camera posture through image datasets or pre-trained models to match a query picture models.

Authorized licenced use limited to: To meet V-design Eye's criteria for real-time, robust, BIRMINGHAM UNIVERSITY. IEEE Xplore was used to get this document on July 25, 2020 at 11:02:09 UTC. There are several limitations. IEEE 2020 (c) 1520-9210 Personal use is allowed, but republication and redistribution require IEEE approval. For further information,go to http://www.ieee.org/publications standards/publications/rights/index.html.This paper has been accepted for publication in a future edition of this journal, but it is still being edited. Before the final publishing, the content may change. Information about the citation:a revolutionary vision-based system with high precision and global localisation In this study, the VB-GPS locating method is proposed. It combines the two types of vision-based positioning In order to get around them, they used the above-mentioned technology.While preserving their benefits, they have limitations. [32] Middelberg et al.a similar attempt to get worldwide, real-time data.

III. METHODOLOGY

Smart glass-with this session, we create an AI-powered smart glass with an integrated camera to aid in picture capturing. These images are delivered to proprietary FRCNN machine learning models, which are installed on smart glasses, for processing. The speech answer is delivered to the Smart glass when the photos are processed, and the user hears it through the glass's built-in speaker. Glass is a set of AI glasses for the visually handicapped and blind.

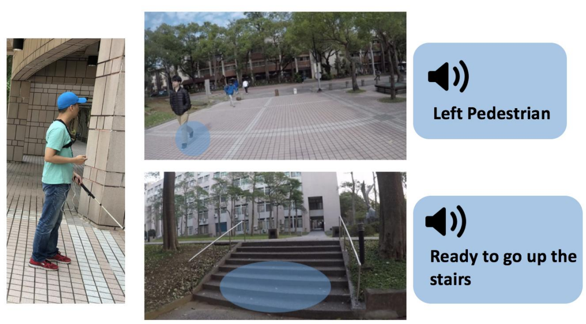
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, June 2022
Impact Factor: 6.252

Face enrolment- this module starts with capturing a few frontal faces of blind people's friends, relatives, and other acquaintances. These templates are then used to evaluate and record the templates for the other poses, such as tilting up and down, moving closer and further, and turning left and right. From the video input, frames are extracted.

Object and face identification-The image is supplied to the face detection module after the Smart Glass Camera captures the object or face image. This module identifies the areas of an image that are most likely to be human. The face detection module uses the Region Proposal Network (RPN) to determine the important features that will be utilised for classification. The face image is given as input to the feature extraction module. The image of the face is then classed as known or unknown.

Prediction-The matching process is carried out in this module using trained classified results and a test Live Camera Captured Classified file. The difference is calculated using Hamming Distance, and the prediction accuracy is displayed as a result. Output of audio. If a data trigger occurs during processing, voice synthesis is utilized to inform the user, for example, by creating the word "halt" if an impediment is encountered. Ramesh, I'd like to say hello.

Performance analysis- using SENSITIVITY, SPECIFICITY, AND ACCURACY of Data in the datasets, we can determine the performance of our system. The datasets are divided into two classes: non-pedestrian (the negative class) and pedestrian (the positive class) (the positive class).True positive (TP), true negative (TN), false negative (FN), and false positive (FP) are used to calculate sensitivity, specificity, and accuracy (FP).The number of positive cases that are classified as positive is denoted by TP.



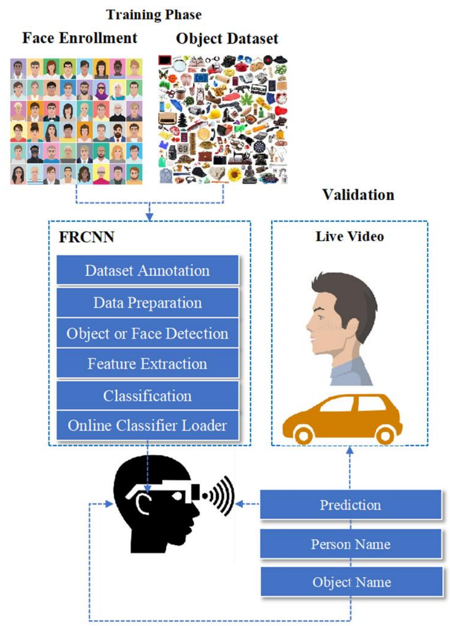
IV. ALOGRITHM

changing the target keyframe's pose from visual SLAM's projected local stance to MBL's estimated global pose The trajectory refinement is applied to the portion in this case.to make the transition between the most recent keyframe and the prior one. The last path is smooth. the procedure, in which A, C, and B denote the position of the most recent. MBL's estimated keyframe, the previous keyframe's position. MBL's estimation, as well as the position of the most recent keyframe computed using the relative poses of previous frames together with visual information SLAM, SLAM, SLAM, SLAM, SLAM Because of this, A and B will frequently disagree. Visual SLAM drift error B will be the final product of the integration procedure.be changed to A, yet there is still a discrepancy between the two.

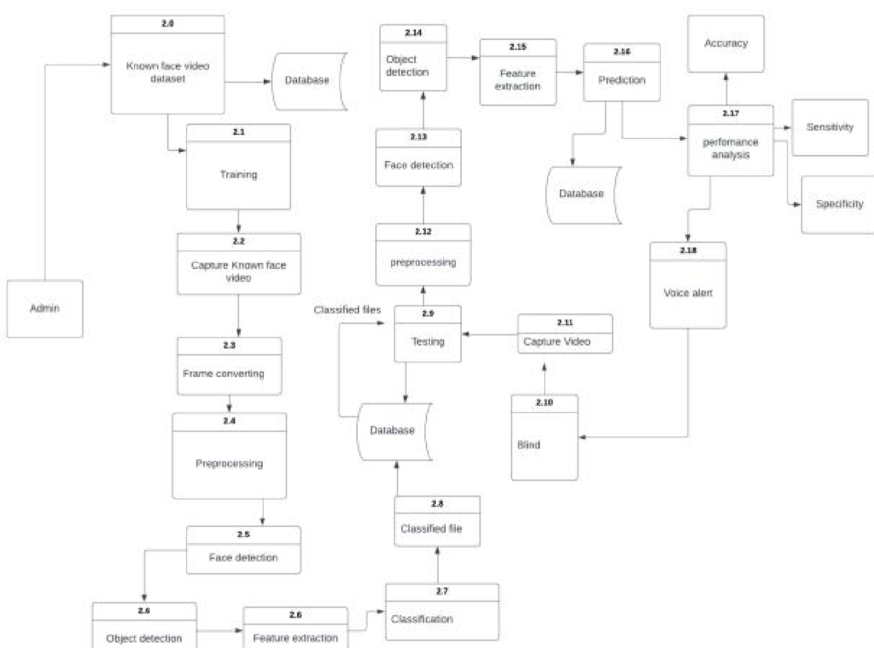
We use trajectory refinement to fill in the gaps and make the transition from C to A as smooth as possible. First, we apply the adaptive scale revision-estimated updated scale, which makes the scale between C and B and the scale between C and

A the same. Then, denote two vectors v1 and v2 as follows: The vectors from C to A and C to B The rotation matrix is then calculated. from version 2 to version 1 Finally, we rotate all of the photos using the rotation matrix R the path that connects C and B It will improve B's position to A's. and smooth out the trajectory The trajectory refinement approach will ensure that the positions of keyframes predicted from MBL are fixed, as well as the relative rotation angles between two consecutive frames obtained via visual SLAM.

V. SYSTEM ARCHITECTURE



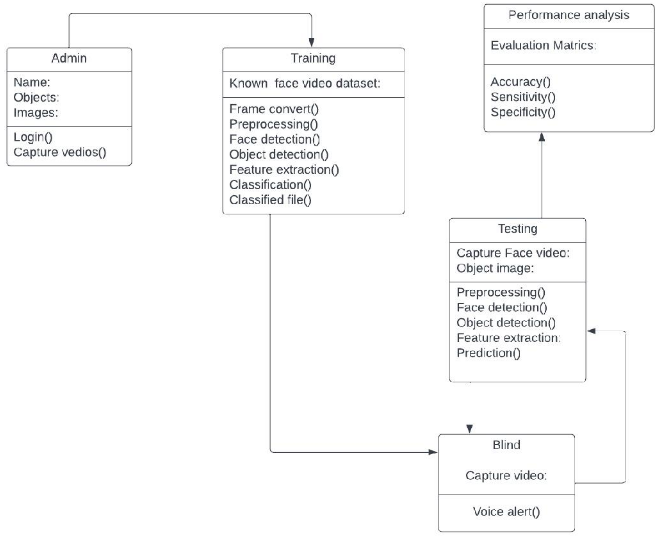
VI. DATAFLOW DIAGRAM



DOI: 10.48175/568

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 1, June 2022

VII. FLOWCHART



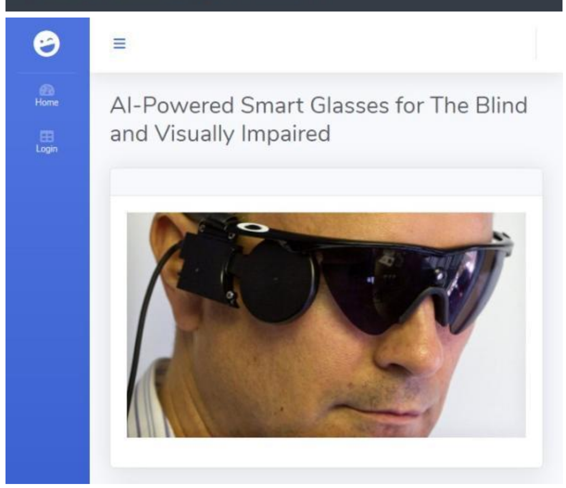
VIII. FINAL PROTOTYPE

```
@app.route('/clear_data',methods=['POST','GET'])
def clear_data():
    ff=open("person.txt","w")
    ff.write("")
    ff.close()

    ff1=open("get_value.txt","w")
    ff1.write("")
    ff1.close()
    return render_template('clear_data.html')

@app.route('/user_view')
def user_view():
    mycursor = mydb.cursor()
    mycursor.execute("SELECT * FROM register")
    result = mycursor.fetchall()
    return render_template('user_view.html', result=result)


@app.route('/logout')
def logout():
    # remove the username from the session if it is there
    #session.pop('username', None)
    return redirect(url_for('index'))
```



IX. RESULT

All three activities were performed successfully by all participants, however the time it took them to do so varied greatly, as shown in Table III. Participants who infrequently participated in orientation and mobility (O&M) training (P5, P6) took twice as long to complete the tasks as the others. Except for P1 and P4, who had amblyopia and could easily keep on track, the participants had trouble changing their course and, as a result, received an increasing amount of signals as the trial progressed. Unpredictable factors also influenced the outcomes of users. For example, P2 had a harder time finishing Task 2 than the others because he encountered more pedestrians; while P5 received more messages on Task 1 than the others, probably due to her comparatively low intelligence. spatial awareness as reported by the individual. P6, who usually walks with her parents, took part in the experiment for the first time on her own, and she received the most messages. Despite their differing levels of vision and expertise walking without assistance, none of the participants travelled the wrong way or missed their three destinations. The sections that follow discuss our observations as well as the study's feedback.

X. CONCLUSION

The item on display is a smart glass that combines the capabilities of a machine vision system with an obstacle detection and recognition sensor. It may be easily published and made accessible to people who are blind or visually impaired. Allow visually impaired people to engage more closely with those around them without worry of their vision becoming clouded and hazy.

REFERENCES

[1]. YOLO-FIRI: Improved YOLOv5 for Infrared Image Object Detection, Year:2021-Shasha Li; Yongjun Li; Yao, Li; Mengjun Li; Xiaorong Xu, Link: https://ieeexplore.ieee.org/document/9576741

[2]. Learning Domain-Invariant Discriminative Features for Heterogeneous Face Recognition, Year:2020Shanmin Yang, Keren, Fu, Xiao Yang, Ye Lin, Jianwei, Zhang, Cheng Peng Link:, https://ieeexplore.ieee.org/document/9262951

[3]. Face Detection Based on Receptive Field Enhanced Multi-Task Cascaded Convolutional Neural Networks, Year:2020 Xiaochao Li, Zhenjie Yang, Hongwei Wu, Link: https://ieeexplore.ieee.org/document/9195457

[4]. A Hybrid Algorithm for Face Detection to Avoid Racial Inequity Due to Dark SkinYear:2021, Muhammad, Syed Sarmad Abbas, Adnan Abid, Saim Rasheed, Link: https://ieeexplore.ieee.org/document/9585604

[5]. Exposing Fake Faces Through Deep Neural Networks Combining Content and Trace Feature ExtractorsYear:2021, Eunji Kim, Sungzoon Cho, Link:https://ieeexplore.ieee.org /document/9531572

[6]. Object Detection in Thermal Spectrum for Advanced Driver-Assistance Systems, Year:2021, Muhammad Ali Farooq, Peter Corcoran, Cosmin Rotariu, Waseem Shariff, Link: https://ieeexplore.ieee.org/document/9618926