IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SUPERVISION (AI Powered Smart Glasses)

[1]Swanand Wirkar, [2]Vaibhav Shinde, [3]Tushar Pawar, [4]Sayee Aher

[1,2,3,4]Student, Dept. of Artificial Intelligence and Data Science, PES's Modern College Of Engineering, Pune, Maharashtra, India

[1] Priyanka Deshpande

[1] Assistant Professor, Dept. of Artificial Intelligence and Data Science, PES's Modern College Of Engineering, Pune, Maharashtra, India

Abstract: SUPERVISION heralds a breakthrough in service technology designed for the visually impaired, combining object recognition, obstacle detection, scenerecognition and optical character recognition (OCR) in smart glasses. Leveraging leading initiatives such as Google Glass and Microsoft Vision AI, SUPERVISION differentiates itself with a comprehensive, AI-driven approach to enhancing seamless environmental awareness and user interaction. Breaking traditional boundaries and using state-of-the-art algorithms, this hands-free platform provides users with better access,freedom and situational awareness. Combining cutting- edge technology, user-friendly design and innovative design, SUPERVISION marks a revolution and heralds afuture unseen by humans, the ability to explore the worldwith confidence, freedom and dignity.

Keywords— Blind Assistance, Natural LanguageProcessing, Depth sensor Technology, IoT and Sensor interaction, Artificial Intelligence.

I. Introduction

SUPERVISION aims to change the lives of visually impaired people by using smart glasses powered by intelligence. Combining advanced sensors and smart algorithms, these glasses can create a digital imageof the user's surroundings. This enables accurate object recognition, making previously daunting taskslike driving on the street or reading the newspaper manageable. In addition to guidance, monitoring can also be effective by identifying sounds such as achild's laughter or the smell of fresh bread. OCR technology increases freedom by allowing users to read printed text [4]. This innovation not only ensures security but also supports community harmony. Think of a confident student at school or a high school student who likes to walk in the evening;all of which promote understanding. Maintenance is more than a tool; It is a beacon of courage that transforms the invisible world into a land of possibilities. Join us to rebuild a safe space and allowblind people to see the world again.

II. AI ENABLED FEATURES

a. Object Recognition

Object recognition in SUPERVISION uses the role of convolutional neural network (CNN) to analyze the data seen by the smart glasses camera [2]. The system processes this information in real time, allowing the user to classify and identify various objects that appearin their view. After recognition, natural language processing (NLP) algorithms are used to createdescriptions and transform visual impressions into instructions that users can understand. The combination of CNN and NLP enables users to receive accurate and timely information about their environment, thus improving their understanding and interaction with their environment.

b. Scene Description

Scene Description uses advanced computer vision technology to provide contextual information about the user's environment. The system creates detailed descriptions by analyzing visual data, including people, activities, and environments. The descriptions cover various topics such as the number of people present, their activities, and the characteristics of the environment. Eventdescriptions enable users to better understand their environment, improve situational awareness, and facilitate decision-making.

c. OCR (Optical Character Recognition)

The OCR function uses complex image processing algorithms to capture text in documents printed with the smart mirror camera [4]. The system separates content from the background through pre-processing techniques, including image enhancement and segmentation. OCR algorithmsthen analyze the content and convert the visible textinto digital format. After conversion, the speech-to-speech system outputs text from the printer, allowing users to hear and understand printed words (for example, reading documents, logos or letters indicating food), making access to it free andeasy.

d. Issue Detection

Issue Detection uses a combination of depth sensors and machine learning algorithms to identify issues affecting the user's path [3]. These depth sensorsmeasure the distance between the user and surrounding objects, and algorithms analyze this spatial information to identify problems. Whendetected, the system will use audio technology togenerate an alarm that will direct users away from danger. This real-time analysis and alert system provides timely warnings of outages, keeping users safe and increasing trust in many areas.

III. CHALLENGES

a) Hardware Optimization

Ensuring that smart glasses strike a balance between functionality and user comfort is a challenge. The need for a lightweight, ergonomicdesign while accommodating components suchas high-resolution cameras and depth sensors requires meticulous engineering to avoid discomfort or discomfort to the user.

b) Data Diversity and Accuracy

Organizing data, objects, and events acrossmultiple sources is critical for powerful machine learning models. However, accurate and representative data involving many real-world situations is difficult to obtain and requires extensive data collection and manipulation efforts to ensure model validity and reliability

c) Real-time processing and Latency

Real-time processing of data and feedback in smart glasses should address computational limitations and reduce latency. Balancing the needs for product knowledge, problem solving, and job descriptions while providing timely user feedback requires better algorithms and unified hardware sharing, which leads to depression.

IV. METHODOLOGY

SUPERVISION consists of powerful hardware withmirrors designed for real-time environmental data collection, processing and feedback. These glasses feature high-resolution cameras with depth sensors thatenable accurate spatial mapping and object detection. Integrated microphones capture ambient sounds to enhance the user experience. This hardware ecosystem seamlessly connects to the onboard processor, providinghigh data rates and low latency. Additionally, the integrated module can facilitate real-time data transfer and instant feedback.

The core of SUPERVISION's work lies in its powerful capabilities in object recognition and machine learning [1]. Carefully selected datasets train state-of-the-art convolutional neural networks (CNN) to accurately identify objects across multiple domains. This optimization is supported by adaptive learning and data augmentation, ensuring efficiency and flexibility. Supporting object recognition, the system combines advanced obstacle detection and spatial analysis. Using depth sensor data, advanced computer vision algorithms analyze spatial configurations and identify potential problems. Real-time voice alerts direct users away from danger, while spatial analysis algorithms provide contextual information to support navigation.

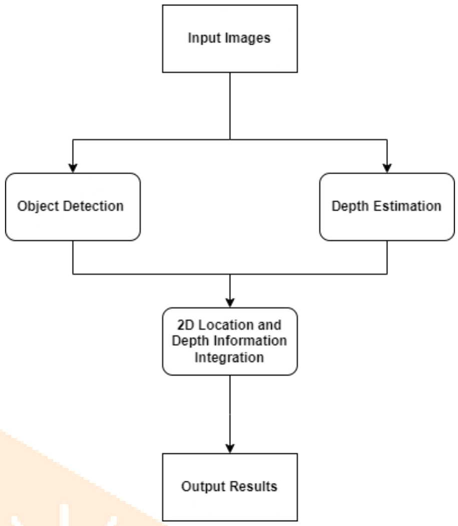

Fig 1: Depth Senor Working Module

Increasing user interaction, SUPERVISION usesdescriptive capabilities that combine computer visionand natural language processing (NLP) [7]. Visual object analysis supported by semantic segmentation is useful for detailed scene information, including spatial dynamics and environmental context. Next-generation NLP improves situational awareness and user engagement by transforming this information into iterative narratives.

In addition, SUPERVISION emphasizes the user'sinteraction with the system through the ability to make noise. The speech recognition algorithm identifies the user's commands and interprets the thoughts correctly. At the same time, the text-to-speech synthesis algorithmindicates the response to the response to clarify the feedback. This voice-centric approach, combined with adaptive interfaces and gesture control, increasesaccessibility and meets user needs and preferences.SUPERVISION combines hardware and softwareinnovations to determine accessibility for the visually impaired through integration, functionality and model for the average user.

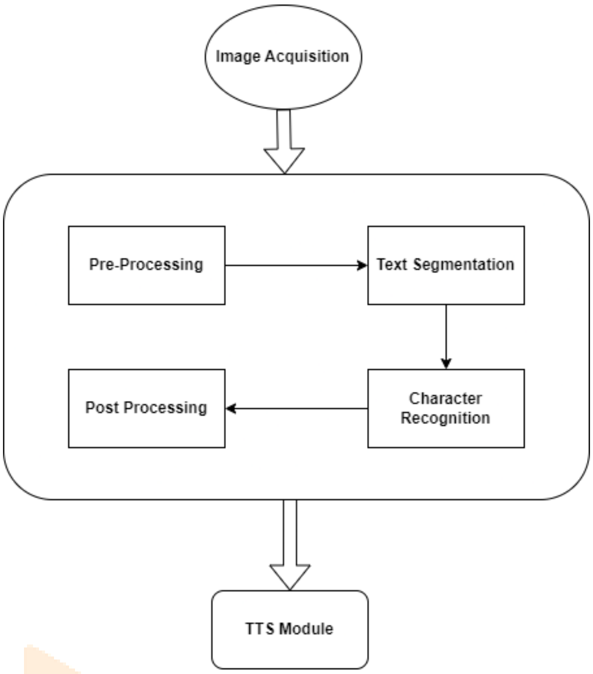

Fig 2: OCR Module Working

As a result, SUPERVISION represents a revolution in service technology for the visually impaired by integrating object recognition, problem detection, nature layer identification and OCR technologies into an integrated mobile platform. SUPERVISION is inspired by major projects such as Google Glass, Microsoft Vision AI and OrCam MyEye 3 and stands out with its integration, artificial intelligence-supported precision and user-friendly design. SUPERVISION strives to determine easy access, independence and quality of life for the visually impaired by solving existing restrictions and expanding the study. With its emphasis on interaction, instant feedback and environmental awareness, SUPERVISION embodies a harmonious combination of technology, innovation and enthusiasm, paving the way for a brighter, better future.

REFERENCES

[1] Mansi Mahendrul, Sanjay Kumar Dubey2(2021),” Real Time Object Detection withAudio Feedback using Yolo vs. Yolo_v3”, from IEEE. 2021 11th International Conference on Cloud Computing, Data Science& Engineering (Confluence) | 978- 1-6654- 1451-7/20/$31.00 ©2021 IEEE | DOI:10.1109/Confluences51648.2021.9377064

[2] Ayoub B, Mustapha A, (2023),”Object Detection Using Deep Learning, CNNs andVision Transformers”, from IEEE, Digital Object Identifier 10.1109/ACCESS.2023.3266093.

[3] S. Ramos, S. Gehrig, P. Pinggera, U. Franke, and C. Rother, “Detecting unexpected obstacles for self-driving cars: Fusing deep learning and geometric modeling.” 2016,arXiv:1612.06573. Accessed: Oct. 21,2019.

[4] X. Chan, H. Ma, J. Wan, B. Li, and T. Xia, ‘‘Multi-view 3D object detection network forautonomous driving,’’ 2016,aeXiv: 1611.07759. Accessed: Oct. 21, 2019.

[5] Nethra Krishnan (2019),” Lidar for Autonomous Driving The principles, challenges, and trends for automotive lidar and perception systems”, from IEEE, DigitalObjectIdentifier 10.1109/MWSCAS.2019.8884887

[6] R. O’Keeffel, S. Gnecchi2, S. Buckley2, C. O’Murchu1, A. Mathewson1 (2018),” Long Range LiDAR Characterisation for Obstacle Detection for use by the visually Impaired and Blind”,from IEEE, 2018 IEEE 68th Electronic Components and Technology Conference, DOI10.1109/ECTC.2018.00084

[7] P.Chitra,V.Balamurugan,M.sumathi,N.Mathan,K.Shrilata,R.N armadha(2021) "Voice Navigation Based guiding Device for Visually Impaired People",from IEEE, International Conference on Artificial Intelligence and SmartSystems (ICAIS-2021), | DOI: 10.1109/ICAIS50930.2021.9395981

[8] Sustainable Development Goals.2020.URL:https://www.un.developmentgoals/

[9] Chucai Yi et al."Finding objects for assistingblindpeople".In:NetworkModellingAnalysisinHealthInformaticsandBioinformatics 2.2(2013), pp. 71-79

[10] G. Shaun K Kane et al. "Freedom to roam: a studyof mobile device adoption and accessibility forpeople with visual and motor disabilities". In:Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility. 2009, pp115– 122

[11] Chandrika Jayant et al. "Supporting blind photography". In: The proceedings of the 13thinternational ACM SIGACCESS conference on Computers and accessibility. 2011, pp.203210.

[12] Menghan Hu et al. "An overview of assistive devicesfor blind and visually impaired people". In:Interna-tional Journal of Robotics andAutomation34.5 (2019),pp. 580- 598.

[13] ChenWorld Health Organization. 2019. URL: https://www.who.int/newsroom/factsheets/detail/blindnessandvisualimnt

[14] PMenghan Hu et al." An overview of assistivedevicesfor blind and visually impaired people". In:Interna-tional Journal of Roboticsand Automation34.5 (2019),pp. 580– 598.

[15] AY. LeCun, Y. B Bengio, and G. Hinton, ''Deep learning,'' Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[16] P. F F Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, ''Object detection with discriminatively trained part- based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, pp. 1627– 1645,Sep. 2010, doi: 10.1109/TPAMI.2009.167.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, vol. 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q.Weinberger, Eds. Red Hook, NY, USA: CurranAssociates,2012,pp.10971105.Accessed:Oct.22 2019.[Online].Available:http://papers.nips.cc/ paper/484imagenetclassificationwithdeepconvolutionalneuralnworks.df

[18] TC. Chan, A. Seff, A. K Kornhauser, and J. Xiao, "DeepDriving: Learning affordance for directperception in autonomous driving,'' in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Santiago, Chile, Dec. 2015, pp. 2722–2730, doi: 10.1109/ICCV.2015.312.

[19] J. Ni, K. Shen, Y. Chen, W. Cao, and S. X. Yang,''An improved deep network-based scene classification method for self- drivingcars,'' IEEE Trans. Instrum. Meas., vol. 71,pp. 1– 14, 2022, doi: 10.1109/TIM.2022.3146923.

[20] https://robu.in/product/rp-lider-a 1 m8-360-degrees-laser-range-finger